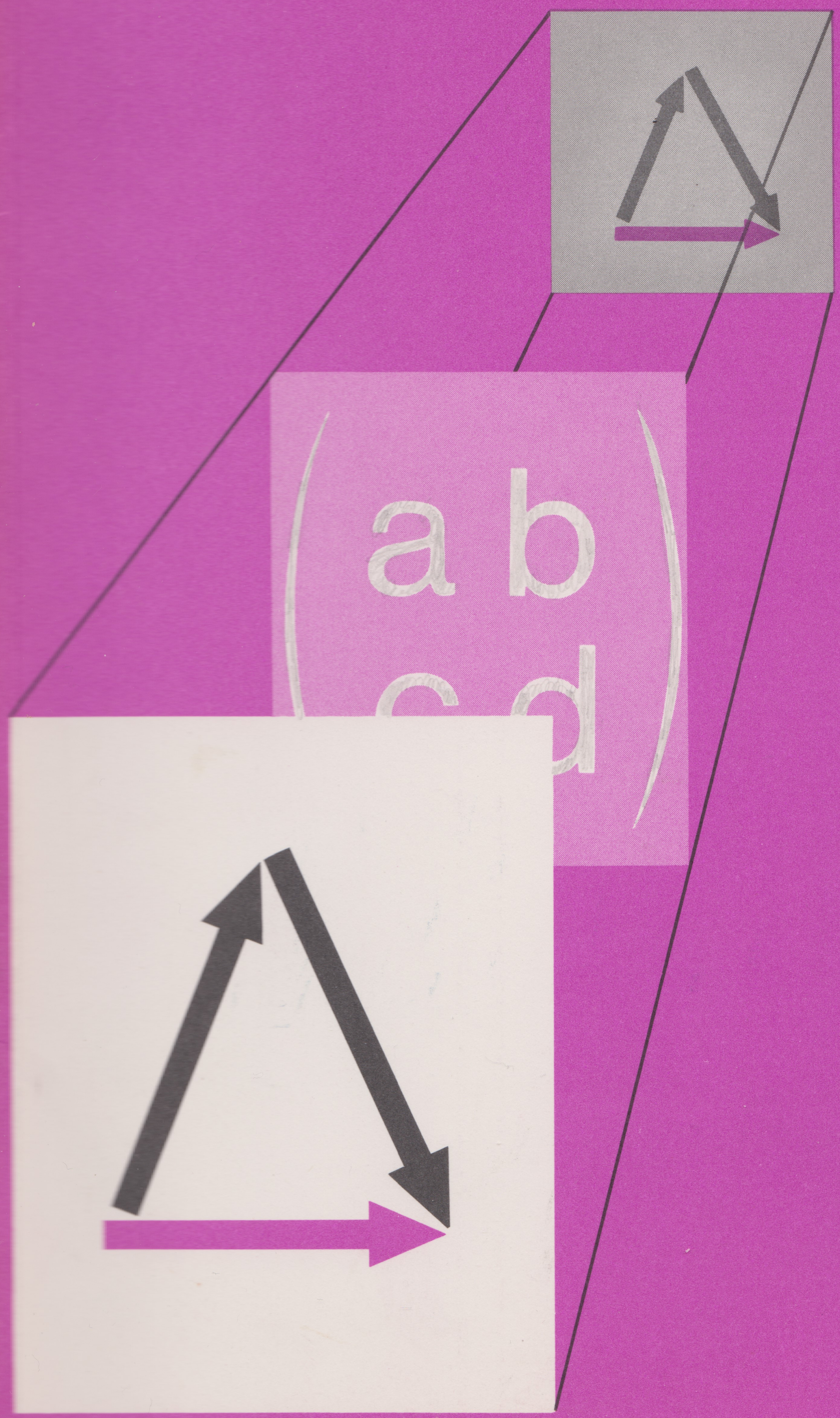
M100 26

THE OPEN UNIVERSITY

Mathematics Foundation Course Unit 26

Linear Algebra III



The Open University

Mathematics Foundation Course Unit 26

LINEAR ALGEBRA III

Prepared by the Mathematics Foundation Course Team

Correspondence Text 26

The Open University Press

Dr. Hans Liebeck acted as consultant for this unit.

Open University courses provide a method of study for independent learners through an integrated teaching system including textual material, radio and television programmes and short residential courses. This text is one of a series that make up the correspondence element of the Mathematics Foundation Course.

The Open University's courses represent a new system of university level education. Much of the teaching material is still in a developmental stage. Courses and course materials are, therefore, kept continually under revision. It is intended to issue regular up-dating notes as and when the need arises, and new editions will be brought out when necessary.

Further information on Open University courses may be obtained from The Admissions Office, The Open University, P.O. Box 48, Bletchley, Buckinghamshire.

The Open University Press
Walton Hall, Bletchley, Bucks

First Published 1971

Printed in Great Britain by
J W Arrowsmith Ltd, Bristol 3

SBN 335 01025 3

Contents

Objectives

The general aim of this unit is to familiarize you with some of the theoretical and practical aspects of the solution of systems of linear equations.

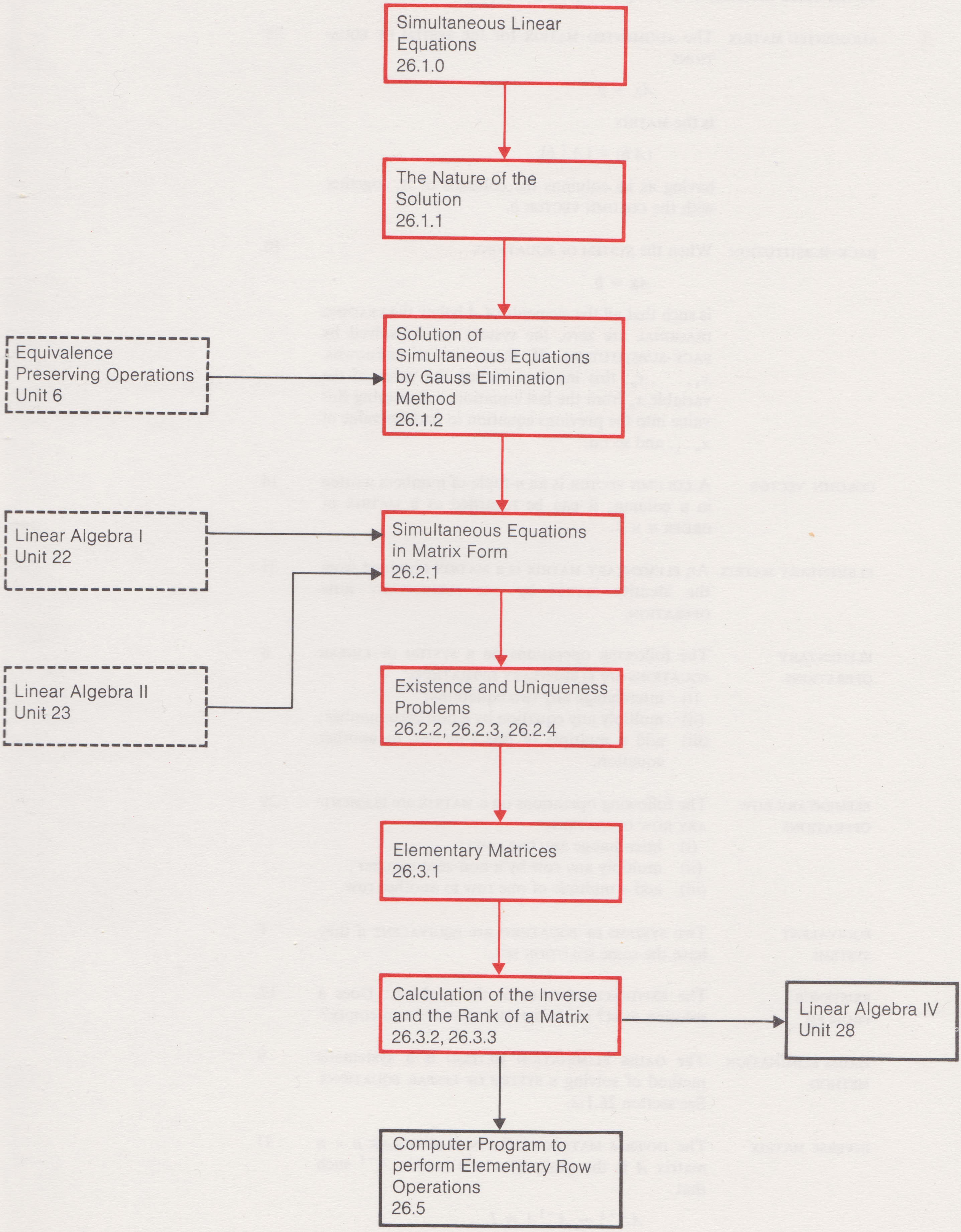
After working through this unit you should be able to:

(i) solve a system of linear equations by the Gauss elimination method;
(ii) discuss the existence and uniqueness problems for systems of linear equations, and give geometrical illustrations in simple cases;
(iii) define elementary row operations on a matrix, and construct elementary matrices corresponding to given elementary operations;
(iv) find the inverse of a given non-singular square matrix;
(v) calculate the rank of a given matrix;
(vi) explain the connection between the rank of a matrix and the existence and uniqueness problems for systems of linear equations.

Note

Before working through this correspondence text, make sure you have read the general introduction to the mathematics course in the Study Guide, as this explains the philosophy underlying the whole course. You should also be familiar with the section which explains how a text is constructed and the meanings attached to the stars and other symbols in the margin, as this will help you to find your way through the text.

Structural Diagram



Glossary

Page

Terms which are defined in this glossary are printed in CAPITALS.

AUGMENTED MATRIX The AUGMENTED MATRIX for the SYSTEM OF EQUATIONS

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

is the MATRIX

$$(A\ \underset{\sim}{b}) = (A \vdots \underset{\sim}{b}),$$

having as its columns the columns of A, together with the COLUMN VECTOR $\underset{\sim}{b}$. 23

BACK-SUBSTITUTION When the SYSTEM OF EQUATIONS 10

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

is such that all the elements of A below the LEADING DIAGONAL are zero, the system can be solved by BACK-SUBSTITUTION. If there are n unknowns, $x_1, \ldots, x_n$, this involves finding the value of the variable x_n from the last equation, substituting this value into the previous equation to find the value of x_{n-1}, and so on.

COLUMN VECTOR A COLUMN VECTOR is an n-tuple of numbers written in a column; it can be regarded as a MATRIX of ORDER $n \times 1$. 14

ELEMENTARY MATRIX An ELEMENTARY MATRIX is a MATRIX obtained from the identity matrix by one ELEMENTARY ROW OPERATION. 31

ELEMENTARY OPERATIONS The following operations on a SYSTEM OF LINEAR EQUATIONS are ELEMENTARY OPERATIONS: 8

(i) interchange any two equations;
(ii) multiply any equation by a non-zero number;
(iii) add a multiple of one equation to another equation.

ELEMENTARY ROW OPERATIONS The following operations on a MATRIX are ELEMENTARY ROW OPERATIONS: 29

(i) interchange any two rows;
(ii) multiply any row by a non-zero number;
(iii) add a multiple of one row to another row.

EQUIVALENT SYSTEMS Two SYSTEMS OF EQUATIONS are EQUIVALENT if they have the same SOLUTION SET. 9

EXISTENCE PROBLEM The EXISTENCE PROBLEM is the problem: Does a solution exist? i.e. Is the SOLUTION SET non-empty? 17

GAUSS ELIMINATION METHOD The GAUSS ELIMINATION METHOD is a systematic method of solving a SYSTEM OF LINEAR EQUATIONS. See section 26.1.2. 9

INVERSE MATRIX The INVERSE MATRIX of the NON-SINGULAR $n \times n$ matrix A is the (unique) $n \times n$ matrix A^{-1} such that 27

$$AA^{-1} = A^{-1}A = I_{n,n}.$$

Page

KERNEL — The KERNEL of a morphism is the set of elements in the domain which map to the zero element in the codomain. 15

LEADING DIAGONAL — The LEADING DIAGONAL of a SQUARE MATRIX is the diagonal from the top left-hand corner to the bottom right-hand corner. 33

LINEAR EQUATION — A LINEAR EQUATION is an equation of the form 2

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \cancel{0}\, b,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n, \cancel{\in R}$. $b \in R$

LINEAR TRANSFORMATION — A LINEAR TRANSFORMATION is a morphism from one vector space to another. 1

MATRIX ($m \times n$) — A MATRIX ($m \times n$) is a rectangular array of numbers, having m rows and n columns. 12

The MATRIX OF COEFFICIENTS of the SYSTEM OF EQUATIONS: 13

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

is

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

NON-SINGULAR MATRIX — A SQUARE MATRIX of ORDER $n \times n$ is NON-SINGULAR if it has RANK n. 27

ORDER OF A MATRIX — The ORDER OF A MATRIX is $m \times n$ if it has m rows and n columns. If $m = n$, it is said to be of order n. 27

RANK OF A MATRIX — The RANK OF A MATRIX is the maximum number of linearly independent column (or row) vectors of the matrix. 21

ROW VECTOR — A ROW VECTOR is an m-tuple of numbers written in a row; it can be regarded as a MATRIX of ORDER $1 \times m$. 14

SINGULAR MATRIX — A SINGULAR MATRIX is a SQUARE MATRIX of ORDER $n \times n$ which has RANK less than n. 27

SOLUTION SET OF A SYSTEM OF EQUATIONS — The SOLUTION SET OF A SYSTEM OF m EQUATIONS in n variables is the subset S of R^n defined by 3

$$S = S_1 \cap S_2 \cap \cdots \cap S_m,$$

where S_i is the solution set of the ith equation (i.e. a set of n-tuples), $i = 1, 2, \cdots, m$.

SQUARE MATRIX — A SQUARE MATRIX is a MATRIX which has the same number of rows as it has columns. 27

		Page
SYSTEM OF SIMULTANEOUS LINEAR EQUATIONS	A SYSTEM OF SIMULTANEOUS LINEAR EQUATIONS is a set of LINEAR EQUATIONS in the same variables. (See also SOLUTION SET.)	2
UNIQUENESS PROBLEM	The UNIQUENESS PROBLEM is the problem: Is the solution unique? i.e. Does the SOLUTION SET have only a single member?	17

Notation

		Page
$\varnothing$	The empty set.	4
A	The matrix A, and also the linear transformation defined by A. For example,	12

$$A = \begin{pmatrix} a_{11} \cdots a_{1n} \\ a_{21} \cdots a_{2n} \\ \vdots \\ a_{m1} \cdots a_{mn} \end{pmatrix}$$

$$A: \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \longmapsto \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}.$$

$\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_n\}$	The basis for R^n, where	19

$$\underset{\sim}{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \underset{\sim}{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \underset{\sim}{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

$\underset{\sim}{a}_i$	The ith column of the matrix A.	19
$r(A)$	The rank of the matrix A.	21
$(A\ \underset{\sim}{b})$	The augmented matrix for the system $A\underset{\sim}{x} = \underset{\sim}{b}$.	23
A^{-1}	The inverse of A, where A is a non-singular matrix; also the inverse mapping, where A is an isomorphism.	27
$I_{n,n}$	The identity matrix of order $n \times n$.	27
$R_1 \longleftrightarrow R_2$	The elementary row operation: interchange rows R_1 and R_2.	29
$R_1 \longmapsto kR_1$	The elementary row operation: multiply row R_1 by the number k.	30
$R_1 \longmapsto R_1 + kR_2$	The elementary row operation: add a multiple, k, of row R_2 to row R_1.	30
E_i	An elementary matrix.	31

Bibliography

S. Lang, *Introduction to Linear Algebra* (Addison-Wesley, 1970).

This book, which we recommended for *Unit 23*, will also prove helpful for this unit. It gives a good discussion on matrices and the solution of systems of linear equations.

D. L. Kreider, R. G. Kuller, D. R. Ostberg and F. W. Perkins, *An Introduction to Linear Analysis* (Addison-Wesley, 1966).

The first two chapters of this book are relevant to the unit. The first chapter deals with vector spaces; the second chapter deals with linear transformations and matrices.

A. E. Coulson, *An Introduction to Matrices* (Longmans, 1965).

This book is a paperback which contains many easy examples and exercises on matrix manipulation.

26.0 INTRODUCTION

26.0

Introduction
* *

In this unit we consider some theoretical aspects of solving systems of simultaneous linear equations. We use both the vector notation introduced in *Unit 22, Linear Algebra I* and the matrix notation introduced in *Unit 23, Linear Algebra II* in an investigation of the existence and uniqueness of solutions to such systems of equations. The aim of the unit is twofold. Firstly, it aims to familiarize you with a *matrix*, both as a computational tool and as a representation of a morphism from one vector space to another. Secondly, it prepares the ground for *Unit 28, Linear Algebra IV*, in which practical methods of solving systems of simultaneous equations will be discussed.

An important theme of this unit is the notion of a *morphism* (also known, in the context of vector spaces, as a *linear transformation*) between two vector spaces. In particular, we shall discuss under what circumstances an inverse morphism exists. Towards the end of this text we discuss a method for finding the inverse morphism (when it exists). This method itself is not very practical, but nevertheless it is important in principle and it is the basis of a number of practical methods.

26.1 SYSTEMS OF LINEAR EQUATIONS

26.1

26.1.0 Introduction

26.1.0

Introduction
* *

We have already considered the idea of simultaneous equations and how they relate to mappings between vector spaces (*Unit 23, Linear Algebra II*). Before we discuss this in greater detail, we shall explain what is meant by a *system of simultaneous linear equations* and a *solution set* of such a system. This will lead us to consider how to find a solution and whether the solution is unique—and indeed whether a solution exists at all.

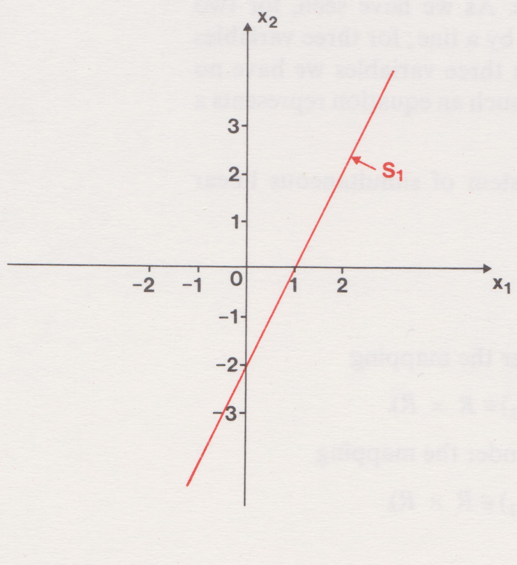
We can consider the equation

$$2x_1 - x_2 = 2,$$

as defining the set of all ordered pairs of real numbers (x_1, x_2) such that $2x_1 - x_2 = 2$; that is,

$$S_1 = \{(x_1, x_2): 2x_1 - x_2 = 2\},$$

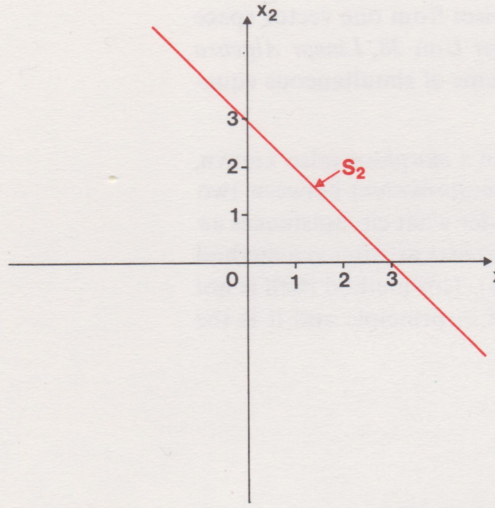
which can be represented diagrammatically as in the following figure:



Similarly, if we consider the equation $x_1 + x_2 = 3$, it defines the set S_2, where

$$S_2 = \{(x_1, x_2): x_1 + x_2 = 3\}.$$

S_2 is represented in the figure below:



We now have two equations:

$$2x_1 - x_2 = 2$$
$$x_1 + x_2 = 3.$$

If we consider these two equations together, they form a *system of simultaneous equations* in the sense that we consider just those pairs (x_1, x_2) which satisfy both equations. Each of the equations has two variables, x_1 and x_2, and each equation defines a straight line in the plane. We therefore say that the equations are *linear*. Thus the system of equations is a set of two simultaneous linear equations in two variables, x_1 and x_2.

In this text we shall deal with systems of linear equations, but we shall generalize to consider m simultaneous equations in n variables (also called *unknowns*). A linear equation in n variables, $x_1, x_2, \ldots, x_n$, is an equation of the form

Definition 1
* * *

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \ldots, a_n$, b are known numbers. As we have seen, for two variables we can represent such an equation by a line; for three variables the representation is a plane; for more than three variables we have no visual representation, but geometers say that such an equation represents a *hyper-plane*.

We now consider the solution set of a system of simultaneous linear equations.

The solution set, S_1, of

$$2x_1 - x_2 = 2$$

is the subset of $R \times R$ which maps to 2 under the mapping

$$f:(x_1, x_2) \longmapsto 2x_1 - x_2 \qquad ((x_1, x_2) \in R \times R).$$

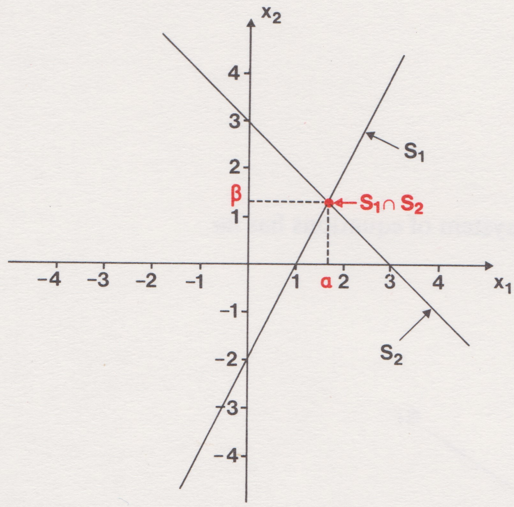
Similarly, S_2 is the subset which maps to 3 under the mapping

$$g:(x_1, x_2) \longmapsto x_1 + x_2 \qquad ((x_1, x_2) \in R \times R).$$

The *solution set of the system* of the two simultaneous equations is the set S of elements of $R \times R$ which map to 2 under f and 3 under g; that is,

$$S = S_1 \cap S_2.$$

In other words, the solution set of the system is the set of ordered pairs $\{(x_1, x_2)\}$ such that each element of this set belongs *both* to the set S_1 and to the set S_2.



We can see from the above diagram that S consists of one element only, namely the point of intersection of the two straight lines. If the co-ordinates of this point are α and β, then

$$S = \{(\alpha, \beta)\}.$$

Definition 2

For a system of m simultaneous linear equations in n variables, $x_1, x_2, \ldots, x_n$, the solution set S is the subset of R^n defined by

$$S = S_1 \cap S_2 \cap \cdots \cap S_m,$$

where S_i is the solution set of the ith equation (a set of n-tuples), $i = 1, 2, \ldots, m$.

26.1.1 The Nature of the Solution

26.1.1

Discussion

* *

Consider the solution set of the system of equations:

$$-2x_1 + 3x_2 = 6$$

$$2x_1 - 3x_2 = 12$$

The graphs of the equations are two parallel lines and hence have no common point, so that if

$$S_1 = \{(x_1, x_2): -2x_1 + 3x_2 = 6\},$$

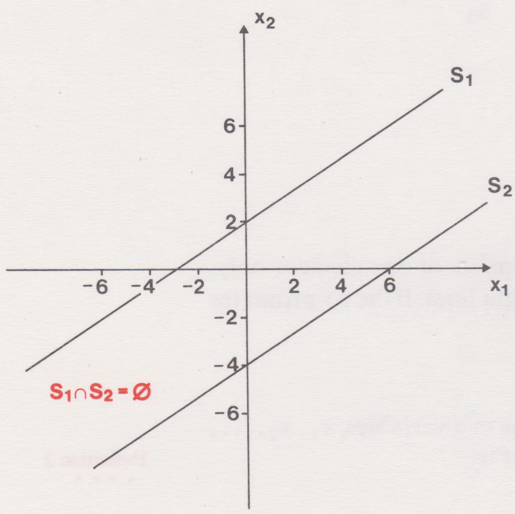
and

$$S_2 = \{(x_1, x_2): 2x_1 - 3x_2 = 12\},$$

then

$$S = S_1 \cap S_2 = \varnothing.$$

The solution set is empty, and we say that the system of equations has *no solution.*



There are examples of systems of linear equations for which the corresponding graphs are non-parallel straight lines and yet the solution sets are empty.

If we consider the equations

$$2x_1 - x_2 = 2$$

$$x_1 + x_2 = 3$$

$$-x_1 + 5x_2 = 5,$$

they define the sets

$$S_1 = \{(x_1, x_2): 2x_1 - x_2 = 2\}$$

$$S_2 = \{(x_1, x_2): x_1 + x_2 = 3\}$$

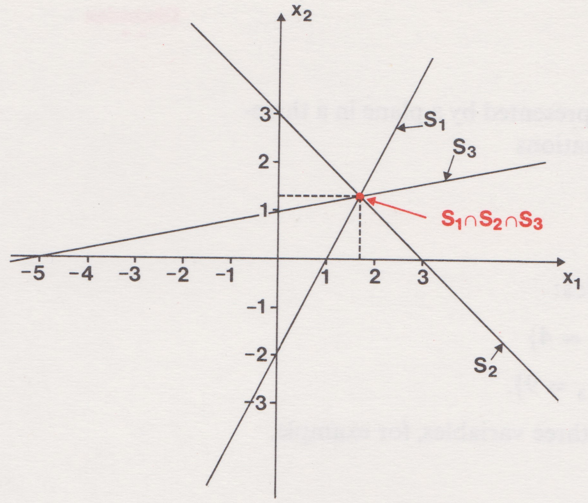
$$S_3 = \{(x_1, x_2): -x_1 + 5x_2 = 5\},$$

and the solution set S of the system is

$$S = S_1 \cap S_2 \cap S_3$$

$$= \{(\tfrac{5}{3}, \tfrac{4}{3})\}.$$

This system has only *one solution.*



On the other hand, the solution set of the system

$$x_1 - 2x_2 = 2$$

$$x_1 - x_2 = -2$$

$$4x_1 + 5x_2 = 20,$$

which defines the three sets

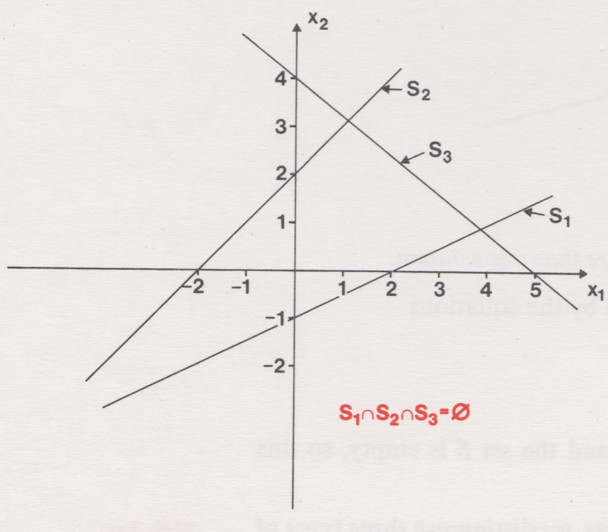
$$S_1 = \{(x_1, x_2): x_1 - 2x_2 = 2\}$$

$$S_2 = \{(x_1, x_2): x_1 - x_2 = -2\}$$

$$S_3 = \{(x_1, x_2): 4x_1 + 5x_2 = 20\},$$

is empty; that is,

$$S = S_1 \cap S_2 \cap S_3 = \varnothing.$$



The figure shows that although any two of the lines intersect, so that

$$S_1 \cap S_2 \neq \varnothing, S_1 \cap S_3 \neq \varnothing \text{ and } S_2 \cap S_3 \neq \varnothing,$$

nevertheless, the three lines do not all intersect at a common point, so that

$$S_1 \cap S_2 \cap S_3 = \varnothing.$$

It follows that the system of equations has *no solution.*

Discussion
★ ★

An equation of the type

$$ax_1 + bx_2 + cx_3 = d$$

(where a, b, c and d are constants) can be represented by a plane in a three-dimensional geometric space. The two equations

$$2x_1 - x_2 + x_3 = 4$$
$$x_1 + 3x_2 - 2x_3 = 9$$

define two sets, S_1 and S_2, of ordered triples:

$$S_1 = \{(x_1, x_2, x_3): 2x_1 - x_2 + x_3 = 4\}$$
$$S_2 = \{(x_1, x_2, x_3): x_1 + 3x_2 - 2x_3 = 9\}.$$

The solution set of an equation involving three variables, for example,

$$2x_1 - x_2 + x_3 = 4$$

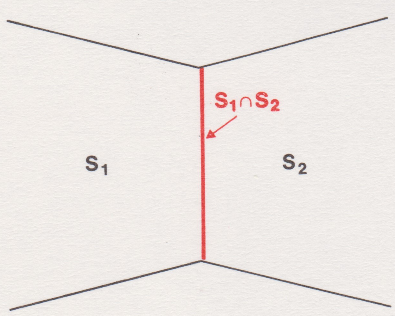
is the set of elements of R^3 (i.e. $R \times R \times R$) which map to 4 under the mapping

$$f:(x_1, x_2, x_3) \longmapsto 2x_1 - x_2 + x_3 \qquad ((x_1, x_2, x_3) \in R^3).$$

Again, the solution set S of the above system is

$$S = S_1 \cap S_2,$$

which is the set of elements (triples of numbers) which are common to both S_1 and S_2. In our example, the two planes meet in a straight line, which therefore defines S.



In this case, the system of equations has *more than one solution.*

On the other hand, the two planes defined by the equations

$$2x_1 - x_2 - x_3 = 4$$
$$4x_1 - 2x_2 - 2x_3 = 7$$

are parallel. The planes do not intersect and the set S is empty, so this system has *no solution.*

Main Text
★ ★ ★

In general, for any system of linear equations, we distinguish three types of solution set:

(i) a solution set which is empty, in which case we say that the system has no solution;

(ii) a solution set which contains just one element (an n-tuple, for a system of equations in n variables), in which case we say that the system has a unique solution;

(iii) a solution set which contains more than one element, in which case we say that the system has more than one solution.

Exercise 1
(3 minutes)

Exercise 1

Examine each of the following systems of linear equations. Distinguish the three types:

A: no solution;
B: unique solution;
C: more than one solution.

Indicate, by a tick in the appropriate box, the class to which each system belongs. (There is no need to solve the equations: a geometrical argument is sufficient, but in case you find it easier to use an algebraic argument, we give some details in the solution.)

(i) $$x_1 - x_2 = 4$$
$$2x_1 + x_2 = 3$$

(ii) $$x_1 + 2x_2 = 1$$
$$-x_1 + 3x_2 = 0$$
$$x_1 + x_2 = 2$$

(iii) $$3x_1 - 4x_2 = -1$$
$$-3x_1 + 4x_2 = 1$$

(iv) $$-x_1 + x_2 = 2$$
$$x_2 + x_3 = 0$$
$$x_1 + x_3 = 1$$

	A	B	C
(i)			
(ii)			
(iii)			
(iv)			

Note
Notice that we ought really to specify how many variables are involved. Thus in (iv), three variables are involved, although each individual equation involves only two. We could make this clear by writing

$$-x_1 + x_2 + 0x_3 = 2,$$

and so on. (Where these equations arise in practice the number of variables is usually clear from the context.) There is a considerable difference between the system in (i) which is taken to involve two variables, and the system

$$x_1 - x_2 + 0x_3 = 4$$
$$2x_1 + x_2 + 0x_3 = 3,$$

so we specify that in the first three parts of this exercise, two variables are involved. ■

Solution 1

Solution 1

	A	B	C
(i)		✓	
(ii)	✓		
(iii)			✓
(iv)	✓		

(i) The solution set is $\{(\frac{7}{3}, -\frac{5}{3})\}$, and consists of a unique element.

(ii) If this system of equations has a solution, it means that the 3 lines represented by these equations intersect at a point. We can find the co-ordinates of the point of intersection of the first two lines and then check whether the third line also passes through it. The first two lines intersect at the point $(\frac{3}{5}, \frac{1}{5})$. Substituting these values for x_1 and x_2 in the third equation gives

$$\tfrac{3}{5} + \tfrac{1}{5} = \tfrac{4}{5} \neq 2.$$

Hence the third line does *not* pass through this point. The solution set of the system is empty.

(iii) On multiplying the second equation by (-1), we see that the two equations have the same solution set. The solution set has many elements; it is:

$$\{(x_1, x_2): 3x_1 - 4x_2 = -1\}.$$

(iv) Adding the first equation to the third, we get

$$x_2 + x_3 = 3$$

and the second equation is

$$x_2 + x_3 = 0.$$

These equations cannot be satisfied simultaneously; there is no solution. ■

26.1.2 Solving Systems of Linear Equations

26.1.2

Main Text
★ ★
Definition 1
★ ★ ★

It may seem fairly obvious to you that the solution set of a system of simultaneous equations is unaffected by the following elementary operations:

(i) interchanging any two equations of the system;
(ii) multiplying every term in an equation by some non-zero constant;
(iii) adding a multiple of one equation to another equation.

We could give some justification for the assertion that these operations do not change the solution set of a system, but without an axiomatic basis for our discussion such an argument would have no real validity. You may, however, like to look again at *Unit 6, Inequalities*, where we discussed the idea of *equivalent* equations or inequalities and *equivalence preserving operations.*

The Method of Gauss Elimination

Main Text

Although the three elementary operations do not change the solution set, they do change the *equations* of the system considered. We obtain a different system of equations which has the same solution set as the original system. Any two linear systems having the same solution set are said to be equivalent systems.

Definition 2

Many methods of solving systems of simultaneous linear equations depend on finding an equivalent system for which it is simple to find the solution set. We shall illustrate this by solving the system of equations:

$$2x_1 + x_2 - x_3 = 3$$

$$6x_1 - x_2 - 9x_3 = 7$$

$$4x_1 + 3x_2 + x_3 = 5$$

We shall use a method known as the Gauss elimination method to solve this system; the method is a formalization of a procedure with which you are probably familiar. We begin by adding appropriate multiples of the first equation to each of the other two so as to eliminate x_1 from the latter two equations. We then add an appropriate multiple of the (new) second equation to the (new) third equation so as to eliminate x_2 from the latter. We then have a system of equations, equivalent to the original system, which can be solved easily.

If we denote the three equations in each of the equivalent systems by R_1, R_2 and R_3 in that order, then we can outline the above sequence of operations in the following diagram.

Action	System	
	System of equations	Variables
	R_1	x_1, x_2, x_3
	R_2	x_1, x_2, x_3
	R_3	x_1, x_2, x_3
Eliminate variable x_1 in R_2 and R_3 by adding multiples of R_1 to R_2 and R_3 to form new R_2 and R_3.		
	Equivalent system of equations	Variables
	R_1	x_1, x_2, x_3
	R_2	x_2, x_3
	R_3	x_2, x_3
Eliminate variable x_2 in R_3 by adding a multiple of R_2 to R_3 to form a new R_3.		
	Equivalent system of equations	Variables
	R_1	x_1, x_2, x_3
	R_2	x_2, x_3
	R_3	x_3

We apply this method to our given system of equations. In order to eliminate the variable x_1 from R_2, we have to multiply R_1 by -3 and add it to R_2; symbolically we have

$$R_2 \longmapsto R_2 + (-3R_1).$$

We then have

$$\begin{aligned}
(R_1)&\quad 2x_1 + x_2 - x_3 = 3\\
(R_2)&\quad -4x_2 - 6x_3 = -2\\
(R_3)&\quad 4x_1 + 3x_2 + x_3 = 5
\end{aligned}$$

To complete the first stage we eliminate x_1 from R_3:

$$R_3 \longmapsto R_3 + (-2R_1),$$

to give

$$\begin{aligned}
(R_1)&\quad 2x_1 + x_2 - x_3 = 3\\
(R_2)&\quad -4x_2 - 6x_3 = -2\\
(R_3)&\quad x_2 + 3x_3 = -1
\end{aligned}$$

We now go to stage 2 and eliminate x_2 from R_3:

$$R_3 \longmapsto R_3 + \tfrac{1}{4}R_2.$$

We obtain

$$\begin{aligned}
(R_1)&\quad 2x_1 + x_2 - x_3 = 3\\
(R_2)&\quad -4x_2 - 6x_3 = -2\\
(R_3)&\quad \tfrac{3}{2}x_3 = -\tfrac{3}{2}
\end{aligned}$$

This system, which is equivalent to the original system, is now in a form which can be easily solved by back-substitution. This means that we obtain the value of x_3 from the last equation, and then substitute this value in the second equation to obtain x_2, and finally we substitute values of x_3 and x_2 in the first equation to obtain x_1.

From R_3: $x_3 = -1$

From R_2: $4x_2 = 2 - 6x_3 \Rightarrow \quad x_2 = 2$

From R_1: $2x_1 = 3 - x_2 + x_3 \Rightarrow x_1 = 0$

The solution set is

$$\{(0, 2, -1)\}.$$

This is the solution set of the given system. You will note that it consists of one element only. Later we shall see just why there are no other elements in the solution set.

The following points should be noted:

(i) The elimination method is systematic. We take no notice of any quick tricks based on the particular numbers in the equations. This is because we want a method which we can discuss theoretically and implement mechanically (on a calculating device).

(ii) The method is deceptively simple. The deception lies in the fact that we have chosen a relatively small system and done our arithmetic exactly. In general, and especially when using a calculating machine, rounding errors will be involved which can, in unfavourable cases, cause serious trouble. Although a discussion of such points would be a useful application of what we have already learnt about error arithmetic in this course, we shall be more interested in theoretical considerations in this unit.

(iii) Slight variations in the method may be necessary. For instance, there may be no x_1 in the first equation. Such variations are dealt with easily in hand calculations, but require care in automatic computing.

(iv) Gauss elimination is an *elimination* method, as opposed to an *iterative* method. In the former, we proceed step by step towards a solution, which is obtained at the *end* of the process. In the latter, we obtain

an estimate of the solution at *each stage* in the process. We do not discuss iterative methods here.

(v) As a numerical method, the method we have described has one deficiency: it has no check built into it. Of course, we can check our solution by direct substitution into the original system, but although this may tell us that we have made an error, it will not tell us where the error occurred. In fact it is quite easy to build a "running check" into the method, which checks each stage in the calculation.

We shall look at some of these points in *Unit 28, Linear Algebra IV*, when we consider one of the practical aspects of solving linear equations: for the present, we shall deal mainly with the theoretical aspects.

Exercise 1

Exercise 1
(3 minutes)

Use the Gauss elimination method to solve the following simultaneous equations:

(i) $3x_1 - 2x_2 = 1$

$4x_1 + x_2 = 3$

(ii) $x_1 - 2x_2 - x_3 = -6$

$x_1 - 2x_2 + 2x_3 = 3$

$-x_1 + x_2 + x_3 = 4$

■

Solution 1

Solution 1

(i) The solution set is $\{(\frac{7}{11}, \frac{5}{11})\}$.

(ii) Eliminating x_1 from R_2 and R_3, we obtain

$$\begin{aligned} x_1 - 2x_2 - x_3 &= -6 \\ 3x_3 &= 9 \\ -x_2 \qquad &= -2 \end{aligned}$$

It is unnecessary to carry on further, since the solutions can be found by back-substitution:

$$x_2 = 2$$

$$x_3 = 3$$

$$x_1 = -6 + x_3 + 2x_2 = 1$$

Hence the solution set is $\{(1, 2, 3)\}$. ■

26.2 MATRICES AND SYSTEMS OF LINEAR EQUATIONS

26.2

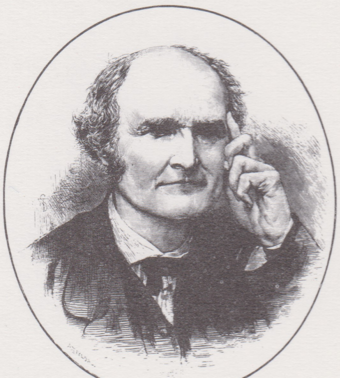
26.2.1 Systems of Linear Equations in Matrix Form

26.2.1

Notation
* * *

We met matrices in *Unit 23, Linear Algebra II*, where they were used as convenient shorthand notation in the context of systems of equations. Now we are going to take the subject a little further. The development of notation is one of the features of mathematics. Notation usually begins as an abbreviation adopted for convenience, but may sometimes lead to significant advances as new concepts evolve around it. Matrices first appeared in about the year 1858, when Cayley introduced the notation:

$$\begin{pmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{21} & a_{22} \cdots a_{2n} \\ \vdots & \vdots \qquad \vdots \\ a_{m1} & a_{m2} \cdots a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$



Arthur Cayley, 1821–1895
(Mansell Collection)

as a shorthand for the system of m simultaneous linear equations in n unknowns, $x_1, x_2, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

You will notice that we use a double suffix notation for the coefficients. At first sight this may seem rather awesome, but it proves to be very useful. The first suffix specifies the equation to which the coefficient belongs, and the second suffix specifies the variable to which the coefficient is attached. Thus the element in the ith row and the jth column of the matrix of coefficients is a_{ij}, and a_{ij} is the coefficient of the variable x_j in the ith equation.

The significant feature of Cayley's shorthand notation is the disentanglement of the array of coefficients from the variables. This allows us to abbreviate still further by writing the system as

$$Ax = b,$$

where A is the matrix of coefficients:

$$A = \begin{pmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{21} & a_{22} \cdots a_{2n} \\ \vdots & \vdots \quad\quad \vdots \\ a_{m1} & a_{m2} \quad\; a_{mn} \end{pmatrix},$$

and x is the (column) matrix of the n variables:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

The right-hand sides of the equations form the (column) matrix b:

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Do you feel that you have "been here before"?

Discussion
* *

By virtue of what it denotes, this notation embodies the definition of premultiplication of a column matrix by another matrix, which we denoted by $\square$ in section 23.2.3 of *Unit 23*. We saw that $\square$ led to the definition of a more general operation of matrix multiplication, which we denoted by $*$. We have already discussed the properties of $*$ in *Unit 23*. (You may like to revise them by re-reading section 23.2.5.) In fact, the only new thing introduced here is the double suffix notation for the matrix elements. There is one other simplification which we have made: in accordance with general practice, we have dropped the symbol $*$ between A and x.

Now that we have written our system of linear equations in matrix form, the solution set is a set of n-element *column matrices*, such that y belongs to the solution set if and only if

$$Ay = b.$$

We know that the solution set may be empty, or consist of one element only, or consist of more than one element.

In section 26.1.2 we solved the system of equations:

$$\begin{aligned} 2x_1 + 1x_2 - 1x_3 &= 3 \\ 6x_1 - 1x_2 - 9x_3 &= 7 \\ 4x_1 + 3x_2 + 1x_3 &= 5, \end{aligned}$$

and found the solution set to be $\{(0, 2, -1)\}$, the set having the one element only. In matrix notation we would write the system of equations as

$$Ax = b$$

where

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 6 & -1 & -9 \\ 4 & 3 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ and } b = \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix}.$$

The solution set consists of one element y, namely

$$y = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}.$$

In section 23.2.1 of *Unit 23, Linear Algebra II*, we showed that every system of m linear equations in n variables defines a *morphism* from the vector space R^n to the vector space R^m, and conversely, any morphism from R^n to R^m can be represented by such a system of linear equations. We saw that if we know the image of a basis of R^n, then we can find the image of *any* element of R^n, expressed as a linear combination of base vectors of R^m. In other words, we can *associate* the matrix A in the equation

$$Ax = b,$$

A being a matrix of order $m \times n$, with a morphism

$$T: \underset{\sim}{x} \longmapsto A\underset{\sim}{x} \qquad (\underset{\sim}{x} \in R^n),$$

which maps R^n to R^m. In this text we assume that we have chosen bases for R^n and R^m, and that we use these bases throughout, so that we *identify* the $m \times n$ matrix A with the morphism from R^n to R^m. (Notice that we use $\underset{\sim}{x}$ for an element of R^n to distinguish it from the particular n-element column matrix x; but we could just as well turn the set of all n-element column matrices into a vector space and regard T(or A) as mapping this space into the space of all m-element column matrices.)

So by introducing an appropriate notation for systems of linear equations, we have a bird's eye view of such systems. Instead of examining in detail *each* equation of the system, we examine the *whole system* and treat it as just one member of all possible such systems. It does not follow that this will necessarily result in any spectacular discoveries, but it may give us a better insight and understanding of these systems.

Discussion * *

There are further advantages in using the matrix notation. For example, we can consider the matrix A to be made up of matrices of smaller order than A itself; such matrices are called *sub-matrices* of A. All these sub-matrices of A fit together to make up the matrix A. As particular examples of this, we can think of the n *columns of elements* of the $(m \times n)$ matrix A as n *column vectors* or as n sub-matrices of order $m \times 1$. Similarly, we can think of the matrix A as one column of m *row vectors*, each row vector having n components (i.e. the components of the corresponding row of the matrix A).

Example 1

Example 1

We can write the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -2 & 7 \end{pmatrix}$$

as

$$A = (b_1 \quad b_2 \quad b_3),$$

where b_1, b_2, b_3 are the matrices

$$b_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 2 \\ -2 \end{pmatrix} \text{ and } b_3 = \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

Also we can write A as

$$A = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

where c_1, c_2 are the matrices

$$c_1 = (1 \quad 2 \quad 3) \text{ and } c_2 = (2 \quad -2 \quad 7). \qquad \blacksquare$$

We call c_1 and c_2 row matrices to distinguish them from the column matrices b_1, b_2 and b_3, which are the columns of the matrix A. Both row and column matrices can be regarded as n- and m-tuples of numbers, and we can regard them as elements of vector spaces isomorphic to R^n and R^m respectively. Hence the names row and column *vectors.*

26.2.2 The Nature of the Solution

26.2.2

Main Text

* *

In the last section we referred back to the work on vector spaces which we covered in *Unit 23, Linear Algebra II*. In this and the following section, we shall again make use of some of the results mentioned there.

Our problem is to solve a system of linear equations, which in matrix form can be written

$$A\underset{\sim}{x} = \underset{\sim}{b},$$

where A is a matrix of order $m \times n$. A defines a *morphism* from R^n to R^m, also denoted by A:

$$A : \underset{\sim}{x} \longmapsto A\underset{\sim}{x} \qquad (\underset{\sim}{x} \in R^n).$$

(Notice that we now underline $\underset{\sim}{x}$ and $\underset{\sim}{b}$, to emphasize that we are considering them as *vectors*, i.e. elements of a vector space.)

We know from section 23.1.4 of *Unit 23*, that we can obtain the required solution set in two parts.

(i) We need just *one element* of the actual solution set, i.e. one vector $\underset{\sim}{x}$ which satisfies the equation

$$A\underset{\sim}{x} = \underset{\sim}{b}. \qquad \text{Equation (1)}$$

(ii) We need the *complete* solution set of the equation

$$A\underset{\sim}{x} = \underset{\sim}{0}, \qquad \text{Equation (2)}$$

where $\underset{\sim}{0}$ is the zero vector (i.e. the column matrix all of whose elements are zero) in R^m. This solution set is the *kernel* of the morphism.

The actual solution set of the original system is then obtained by adding the solution in (i) to each solution in (ii).

We shall verify this result again in our particular circumstances to remind you of the argument.

Suppose that $\underset{\sim}{x}_p$ is a solution of Equation (1), and $\underset{\sim}{x}_k$ is a member of the kernel, K (i.e. a solution of Equation (2)); then

$$\begin{aligned} A(\underset{\sim}{x}_p + \underset{\sim}{x}_k) &= A\underset{\sim}{x}_p + A\underset{\sim}{x}_k && (A \text{ is a morphism}) \\ &= \underset{\sim}{b} + \underset{\sim}{0} && (\text{hypothesis}) \\ &= \underset{\sim}{b} && (\text{definition of } \underset{\sim}{0}) \end{aligned}$$

This shows that $(\underset{\sim}{x}_p + \underset{\sim}{x}_k)$ is an element of the solution set of Equation (1).

Also, *every* element of the solution set of Equation (1) can be written in the form $\underset{\sim}{x}_p + \underset{\sim}{x}_k$. For if we let $\underset{\sim}{x}_1$ be any element of that solution set, then

$$\begin{aligned} A(\underset{\sim}{x}_1 - \underset{\sim}{x}_p) &= A\underset{\sim}{x}_1 - A\underset{\sim}{x}_p \\ &= \underset{\sim}{b} - \underset{\sim}{b} \\ &= \underset{\sim}{0} \end{aligned}$$

It follows that the vector $(\underset{\sim}{x}_1 - \underset{\sim}{x}_p)$ is a solution of Equation (2), i.e. an element of K, so that we can write

$$\underset{\sim}{x}_1 - \underset{\sim}{x}_p = \underset{\sim}{x}_k,$$

where $\underset{\sim}{x}_k \in K$, and so

$$\underset{\sim}{x}_1 = \underset{\sim}{x}_p + \underset{\sim}{x}_k.$$

Now $\underset{\sim}{x}_k$ is unique, by Axiom 1 for a vector space, so it follows that there is a one-one correspondence between the solution set and the kernel. Thus, the solution set is

$$\{\underset{\sim}{x}_i : \underset{\sim}{x}_i = \underset{\sim}{x}_p + \underset{\sim}{x}_k, \underset{\sim}{x}_k \in K\}.$$

$\underset{\sim}{x}_p$ is called a particular solution of $A\underset{\sim}{x} = \underset{\sim}{b}$.

Definition 1
* *

Example 1

Example 1

In *Unit 23, Linear Algebra II* (section 23.1.4), we considered the system of equations

$$\begin{aligned} 2x + 3y + (-1)z &= 1 \\ 1x + 1y + (-1)z &= 2 \end{aligned}$$

In matrix form it becomes

$$\begin{pmatrix} 2 & 3 & -1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

We found that a particular solution is

$$\underset{\sim}{x}_p = \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix},$$

and the kernel is the set

$$K = \left\{ \begin{pmatrix} 2r \\ -r \\ r \end{pmatrix} : r \in R \right\}.$$

The solution set of the system in matrix or column vector form is therefore:

$$\left\{ \begin{pmatrix} 5 + 2r \\ -3 - r \\ 0 + r \end{pmatrix} : r \in R \right\}.$$

■

Exercise 1

Exercise 1
(3 minutes)

Find the matrix form of the solution set of the system

$$x + \ \ y + z = 4$$

$$2x + 2y - z = 5$$

■

Existence and Uniqueness Problems

Discussion
* *

In section 26.1.1 we noticed that some systems have an empty solution set, whereas others have a non-empty solution set. One of the important problems in the theory of linear equations is to determine conditions under which a system of equations has a non-empty solution set: this problem is called the existence problem.

Definition 2
* * *

Once it has been decided whether or not a solution exists, it is useful to know the conditions under which the solution set contains just *one* element: this problem is known as the uniqueness problem.

Definition 3
* * *

We have seen that the solution set of the equation

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

is

$$\{\underset{\sim}{x} : \underset{\sim}{x} = \underset{\sim}{x}_p + \text{any element of the kernel}\}.$$

Thus, if we can find a particular solution $\underset{\sim}{x}_p$ and determine the nature of the kernel of the mapping, the uniqueness problem is solved.

If the kernel consists of one element only (which must then be the zero element), then the solution set consists of one element only, $\underset{\sim}{x}_p$, and so we have established uniqueness. If, on the other hand, the kernel consists of more than one element, then the system has more than one solution.

As far as existence is concerned, we are assured of a solution if $\underset{\sim}{b}$ is in the image set of the mapping defined by A, for in that case there must be some vector $\underset{\sim}{x}$ which maps to $\underset{\sim}{b}$.

Thus the existence problem is essentially the problem of finding a test by which we can look at A and $\underset{\sim}{b}$ and determine whether $\underset{\sim}{b}$ is in the image set. If the solution set is non-empty, the uniqueness problem will be solved if we can find a way of determining whether the kernel of the mapping defined by A contains more than one element. In the following sections we shall consider these problems in more detail.

Solution 1

Solution 1

If we try to find a particular solution by putting $z = 0$, we obtain

$$x + y = 4$$
$$2x + 2y = 5.$$

We see that there is *no* particular solution of the original system of the form

$$\begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix}.$$

We therefore try putting another variable, y say, equal to zero; we obtain the system

$$x + z = 4$$
$$2x - z = 5,$$

which has the unique solution $x = 3$ and $z = 1$. So

$$\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

is a particular solution of the original system. We now wish to find the kernel i.e. to solve the system

$$x + y + z = 0$$
$$2x + 2y - z = 0.$$

The solution set of this system is

$$\left\{ \begin{pmatrix} r \\ -r \\ 0 \end{pmatrix} : r \in R \right\}.$$

Hence the required solution set is

$$\left\{ \begin{pmatrix} 3 + r \\ -r \\ 1 \end{pmatrix} : r \in R \right\}.$$

■

26.2.3 The Existence Problem

26.2.3

Main Text
* * *

We consider a system of m simultaneous equations in n unknowns, $x_1, x_2, \ldots, x_n$, written in matrix form as

$$A\underset{\sim}{x} = \underset{\sim}{b},$$

where

$$\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \underset{\sim}{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

and A is a matrix of order $m \times n$:

$$A = \begin{pmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{22} \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{mn} \end{pmatrix}$$

In section 26.2.1 we remarked that we can consider a matrix to be made up of vectors. For example, we could write

$$\underset{\sim}{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \underset{\sim}{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \underset{\sim}{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

If we choose the following basis for R^n:

$$\underset{\sim}{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \underset{\sim}{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \underset{\sim}{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

then

$$\underset{\sim}{x} = x_1\underset{\sim}{e}_1 + x_2\underset{\sim}{e}_2 + \cdots + x_n\underset{\sim}{e}_n$$

and

$$\begin{aligned} A\underset{\sim}{x} &= x_1 A\underset{\sim}{e}_1 + x_2 A\underset{\sim}{e}_2 + \cdots + x_n A\underset{\sim}{e}_n \\ &= x_1\underset{\sim}{a}_1 + x_2\underset{\sim}{a}_2 + \cdots + x_n\underset{\sim}{a}_n. \end{aligned}$$

Example 1

Example 1

If we write the system of equations

$$2x_1 + 3x_2 = 1$$

$$5x_1 + 1x_2 = 3$$

in matrix form:

$$\begin{pmatrix} 2 & 3 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

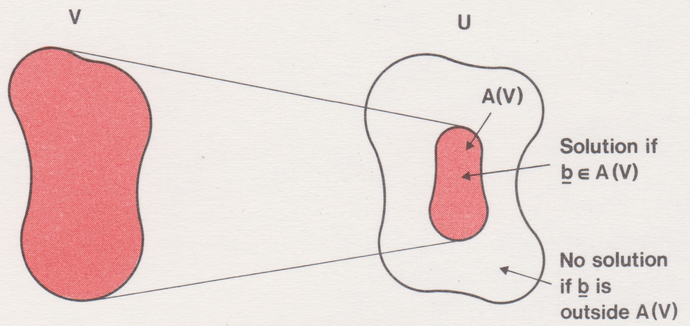
then we can write this as

$$x_1\begin{pmatrix} 2 \\ 5 \end{pmatrix} + x_2\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

This illustrates the general formula given above. ■

Main Text
* * *

Now let us return to our problem of the existence of a non-empty solution set. $A\underset{\sim}{x}$ is a vector in the image space $A(R^n)$. Since $A\underset{\sim}{x} = x_1\underset{\sim}{a}_1 + x_2\underset{\sim}{a}_2 + \cdots + x_n\underset{\sim}{a}_n$, this means that any vector in the image space is a linear combination of the vectors $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_n$. For a solution to exist, $\underset{\sim}{b}$ must be in $A(R^n)$, and so $\underset{\sim}{b}$ must be a linear combination of $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_n$.



(Here $V = R^n$.)

So we have what appears to be a simple test which we can apply to find whether a system has a solution. Unfortunately the test is not always simple to apply in practice.

Theorem
* * *

The Test

If $\underset{\sim}{b}$ is a linear combination of the vectors $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_n$, then the system of linear equations represented by $A\underset{\sim}{x} = \underset{\sim}{b}$ has a solution. Otherwise the system has no solution (that is, the solution set is empty).

Example 2

Example 2

Consider the system of equations represented by

$$\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 6 \end{pmatrix}.$$

We must decide whether the vector $\underset{\sim}{b} = \begin{pmatrix} 10 \\ 6 \end{pmatrix}$ is a linear combination of the 2 vectors

$$\underset{\sim}{a}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \underset{\sim}{a}_2 = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

In fact, we notice that

$$\underset{\sim}{a}_2 = 2\underset{\sim}{a}_1, \text{since} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2 \times \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

so that effectively our problem reduces to deciding whether $\underset{\sim}{b}$ is a (scalar) multiple of $\underset{\sim}{a}_1$. In other words, is there a number α such that

$$\underset{\sim}{b} = \alpha\underset{\sim}{a}_1,$$

i.e.

$$\begin{pmatrix} 10 \\ 6 \end{pmatrix} = \alpha\begin{pmatrix} 2 \\ 1 \end{pmatrix}?$$

The equation above holds only if the equations

$$10 = 2\alpha$$
$$6 = \alpha$$

can be solved simultaneously. Clearly they cannot. It follows that $\underset{\sim}{b}$ is *not* a linear combination of $\underset{\sim}{a}_1$ and $\underset{\sim}{a}_2$, so that there is *no solution* to the original system of equations. ■

Exercise 1

Exercise 1
(4 minutes)

Using the theorem given in the text, determine which of the following systems of equations have at least one solution.

(i) $x_1 + 2x_2 = 5$

$\quad x_1 + x_2 = 3$

(ii) $x_1 + 2x_2 = 3$

$\quad 3x_1 + 6x_2 = 9$

(iii) $\frac{1}{2}x_1 + \frac{1}{3}x_2 = 2$

$\quad \frac{3}{2}x_1 + x_2 = 8$

(iv) $0.2x_1 + 0.3x_2 + 0.1x_3 = 1.1$

$\quad 0.6x_1 + 0.9x_2 + 0.3x_3 = 2.2.$ ■

Main Text
* * *

In the examples considered so far, the test to determine whether the system $A\underset{\sim}{x} = \underset{\sim}{b}$ has a solution was easy to apply. In fact, for a "small" system we do not need a test; we can establish whether or not a solution exists by trying to solve the equations directly. But for "meatier" systems a test is required. However, in such a case, to find whether $\underset{\sim}{b}$ is a linear combination of the columns of A, we would need to solve a system of simultaneous equations, which could be as complicated as the original system! Clearly, we must modify the test to make it easier to apply. We do this by introducing the concept of the *rank* of a matrix.

Definition 1
* * *

Notation 1
* * *

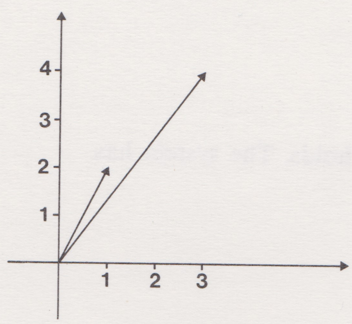
Given the matrix $A = (\underset{\sim}{a}_1\ \underset{\sim}{a}_2 \cdots \underset{\sim}{a}_n)$, the rank of A is the maximum number of linearly independent vectors from the set $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n\}$; it is denoted by $r(A)$.

Example 3

Example 3

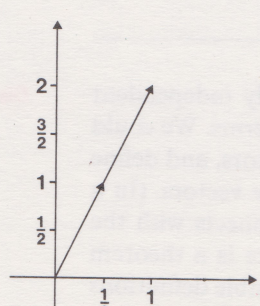
The matrix $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ has rank 2, since the vectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ are linearly independent:

$$\alpha_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \alpha_1 = \alpha_2 = 0.$$



The matrix $\begin{pmatrix} 1 & \frac{1}{2} \\ 2 & 1 \end{pmatrix}$ has rank 1, since the vectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}$ are linearly dependent:

$$\frac{1}{2}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + (-1)\begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$



■ *(continued on page 22)*

Solution 1

Solution 1

(i) The system $A\underset{\sim}{x} = \underset{\sim}{b}$ is

$$\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}5\\3\end{pmatrix}.$$

By inspection,

$$\underset{\sim}{b} = \begin{pmatrix}5\\3\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix} + 2\begin{pmatrix}2\\1\end{pmatrix},$$

so $\underset{\sim}{b}$ is a linear combination of the column vectors of the matrix of coefficients A. It follows that the system has at least one solution.

(ii) The system $A\underset{\sim}{x} = \underset{\sim}{b}$ is

$$\begin{pmatrix}1 & 2\\3 & 6\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}3\\9\end{pmatrix}.$$

In this case,

$$\underset{\sim}{b} = \begin{pmatrix}3\\9\end{pmatrix} = \begin{pmatrix}1\\3\end{pmatrix} + \begin{pmatrix}2\\6\end{pmatrix}.$$

It follows that the system has at least one solution.

(iii) The system $A\underset{\sim}{x} = \underset{\sim}{b}$ is

$$\begin{pmatrix}\frac{1}{2} & \frac{1}{3}\\\frac{3}{2} & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}2\\8\end{pmatrix}.$$

The two column vectors of A are linearly dependent:

$$\begin{pmatrix}\frac{1}{2}\\\frac{3}{2}\end{pmatrix} = \frac{3}{2} \times \begin{pmatrix}\frac{1}{3}\\1\end{pmatrix}$$

So the system will only have a solution if the vector $\underset{\sim}{b}$ is a multiple of (say) the vector $\begin{pmatrix}\frac{1}{3}\\1\end{pmatrix}$, i.e. if

$$\begin{pmatrix}2\\8\end{pmatrix} = \alpha\begin{pmatrix}\frac{1}{3}\\1\end{pmatrix}, \text{ where } \alpha \in R.$$

There is no α such that the above equation holds. The system has no solution.

(iv) The system $A\underset{\sim}{x} = \underset{\sim}{b}$ is

$$\begin{pmatrix}0.2 & 0.3 & 0.1\\0.6 & 0.9 & 0.3\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}1.1\\2.2\end{pmatrix}.$$

The three column vectors of A, $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \underset{\sim}{a}_3$, are linearly dependent; in fact,

$$\underset{\sim}{a}_1 = 2 \times \underset{\sim}{a}_3 \quad \text{and} \quad \underset{\sim}{a}_2 = 3 \times \underset{\sim}{a}_3.$$

But the vector $\underset{\sim}{b}$ is not a multiple of $\underset{\sim}{a}_3$. It follows that the system has no solution. ■

(*continued from page 21*)

Discussion

We have defined rank in terms of the number of linearly independent column vectors, because our discussion has been in these terms. We could equally well consider the matrix to be made up of *row* vectors, and define rank in terms of the number of linearly independent row vectors. (In a sense, this would be quite natural, since intuitively it connects with the number of "independent" equations in the system.) There is a theorem (which we shall not prove) which states that these two possible definitions of the rank of a matrix are equivalent, i.e. give the same value for the rank.

Exercise 2

Exercise 2
(4 minutes)

Find the rank of each of the following matrices.

(i) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$

(ii) $\begin{pmatrix} 2 & 1 & -1 \\ 4 & 1 & 1 \\ 6 & 1 & 3 \end{pmatrix}$

(iii) $\begin{pmatrix} 6 & 3 & 9 \\ 2 & 1 & 3 \\ 4 & 2 & 6 \end{pmatrix}$

■

Discussion
★ ★

We now reconsider our test for the existence of a non-empty solution set in terms of the rank concept.

We have seen that for the system

$$A\underset{\sim}{x} = \underset{\sim}{b}, \text{ where } A = (\underset{\sim}{a}_1 \cdots \underset{\sim}{a}_n)$$

to have a solution, the vector $\underset{\sim}{b}$ must be a linear combination of the vectors

$$\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n.$$

Another way of saying this is that the number of linearly independent vectors in the two sets $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n, \underset{\sim}{b}\}$ and $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n\}$ must be the same.

In terms of matrices, the last remark means that the rank of the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{21} & a_{22} \cdots a_{2n} \\ \vdots & \vdots \quad\;\; \vdots \\ a_{m1} & a_{m2} \cdots a_{mn} \end{pmatrix}$$

Definition 2
★ ★ ★

and the rank of the augmented matrix (the matrix A with the extra column $\underset{\sim}{b}$):

$$(A\ \underset{\sim}{b}) = \begin{pmatrix} a_{11} & a_{12} \cdots a_{1n} & b_1 \\ a_{21} & a_{22} \cdots a_{2n} & b_2 \\ \vdots & \vdots \quad\;\; \vdots & \vdots \\ a_{m1} & a_{m2} \cdots a_{mn} & b_m \end{pmatrix}$$

must be the same.

We can now phrase the test for the existence of a non-empty solution set in the following form:

Theorem
★ ★ ★

The system of linear equations represented by

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

has a non-empty solution set if the matrix A and the augmented matrix $(A\ \underset{\sim}{b})$ have the same rank.

Unfortunately, with our present techniques, to find whether $r(A)$ is the same as $r(A\ \underset{\sim}{b})$ may still necessitate solving a system of equations which is as large as the original system we wish to solve. We shall discuss another possible method for determining the rank of a matrix (which is not dependent on solving simultaneous equations) in section 26.3.3.

Solution 2 **Solution 2**

(i)
$$\alpha_1\begin{pmatrix}1\\0\\1\end{pmatrix} + \alpha_2\begin{pmatrix}1\\0\\0\end{pmatrix} + \alpha_3\begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$

$$\Rightarrow \qquad \alpha_1 + \alpha_2 = 0$$
$$\alpha_3 = 0$$
$$\alpha_1 + \alpha_3 = 0$$
$$\Rightarrow \qquad \alpha_1 = \alpha_2 = \alpha_3 = 0.$$

This shows that the vectors $\underline{a}_1$, $\underline{a}_2$ and $\underline{a}_3$ are linearly independent, and hence $r(A) = 3$.

(ii) $\underline{a}_1 = \begin{pmatrix}2\\4\\6\end{pmatrix}$, $\underline{a}_2 = \begin{pmatrix}1\\1\\1\end{pmatrix}$, $\underline{a}_3 = \begin{pmatrix}-1\\1\\3\end{pmatrix}$. You can verify that $3\underline{a}_2 + \underline{a}_3 = \underline{a}_1$.

It follows that the maximum number of linearly independent column vectors is less than 3. Looking at the expression

$$\alpha_2\underline{a}_2 + \alpha_3\underline{a}_3 = \underline{0},$$

we see that this is equivalent to

$$\alpha_2 - \alpha_3 = 0$$
$$\alpha_2 + \alpha_3 = 0$$
$$\alpha_2 + 3\alpha_3 = 0$$

The first two equations give $\alpha_2 = 0$, and it then follows that $\alpha_3 = 0$. Thus $\underline{a}_2$ and $\underline{a}_3$ are linearly independent, so $r(A) = 2$.

(iii) $\underline{a}_1 = 2\underline{a}_2, \underline{a}_3 = 3\underline{a}_2$.

In this case the maximum number of linearly independent column vectors is 1, so that $r(A) = 1$. ■

26.2.4 The Uniqueness Problem

26.2.4

Main Text
* * *

Having discussed some theoretical aspects of the existence problem, we now turn our attention to the uniqueness problem.

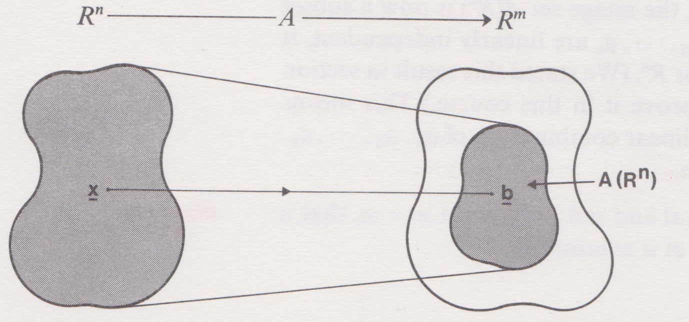
In general, corresponding to the equation

$$A\underset{\sim}{x} = \underset{\sim}{b},$$

where A has order $m \times n$, we have a mapping

$$A : \underset{\sim}{x} \longmapsto A\underset{\sim}{x}$$

of R^n to R^m. We know that, for the equation $A\underset{\sim}{x} = \underset{\sim}{b}$ to have a solution, $\underset{\sim}{b}$ must be a linear combination of the vectors $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n$. So we can determine the image set $A(R^n) \subseteq R^m$, as the set spanned by $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n$.



The dimension of this image set is, therefore, the maximum number of linearly independent vectors among the $\underset{\sim}{a}$'s. But this is what we have defined to be the rank of A, so we have

$$r(A) = \text{dimension of } A(R^n).$$

We can get some interesting results by recalling the dimension theorem of *Unit 23, Linear Algebra II*, section 23.1.6. This stated, in terms of our context, that

$$\text{dimension of } A(R^n) = \text{dimension of } R^n - \text{dimension of kernel},$$

i.e.

$$r(A) = n - \text{dimension of kernel}.$$

Now let us suppose that the solution of $A\underset{\sim}{x} = \underset{\sim}{b}$ is unique. We know in general that the set of all solutions is given by

$$\{\underset{\sim}{x} : \underset{\sim}{x} = \underset{\sim}{x}_p + \underset{\sim}{x}_k, \underset{\sim}{x}_k \in K\},$$

where $\underset{\sim}{x}_p$ is a particular solution and K is the kernel of the mapping A. But if, as we suppose, the solution is unique, then there is only one solution, so the kernel contains just the one element, the zero vector.* This means that the dimension of the kernel is zero, i.e. our result above now becomes

$$r(A) = n.$$

Thus a *necessary* condition for the solution of $A\underset{\sim}{x} = \underset{\sim}{b}$ to be unique is that the rank of $A = n$, the number of variables in the original equations. This condition is obviously also *sufficient*, i.e. it guarantees uniqueness. Because if $r(A) = n$, then the dimension theorem tells us that the dimension of the kernel is zero, and therefore the kernel is the zero vector space. It follows that, if $A\underset{\sim}{x} = \underset{\sim}{b}$ has a solution, then it is unique. So we have:

Theorem
* * *

If a system of m linear equations in n unknowns represented by

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

* Since $\alpha\underset{\sim}{0} = \underset{\sim}{0}$, where $\underset{\sim}{0}$ is the zero vector in a vector space, and $\alpha \neq 0$, $\{\underset{\sim}{0}\}$ is linearly dependent. The dimension of $\{\underset{\sim}{0}\}$ is defined to be zero. It is the only vector space of dimension zero.

has a solution, then the solution is unique if and only if

$$r(A) = n.$$

We can improve on this result if we simplify our case. For, if we assume that $n = m$, i.e. the number of equations is equal to the number of variables, then we can include existence with uniqueness. That is,

Theorem ★ ★ ★

The system of n equations in n unknowns represented by

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

has a unique solution if and only if

$$r(A) = n.$$

We have already shown that, if a solution exists, it is unique, and to prove existence is not difficult. Since $n = m$, the image set $A(R^n)$ is now a subset of R^n. Since the column vectors $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n$ are linearly independent, it can be shown that they form a basis for R^n. (We stated this result in section 22.2.3 of *Unit 22*, but we shall not prove it in this course.) This means that *any* $\underset{\sim}{b} \in R^n$ can be expressed as a linear combination of $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n$. It follows that $A\underset{\sim}{x} = \underset{\sim}{b}$ has a solution.

Discussion ★ ★

In section 26.2.3 we showed, in general and not only when $n = m$, that a *sufficient* condition for the existence of a solution of

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

is that

$$r(A) = r(A\ \underset{\sim}{b}).$$

We have now shown that when $n = m$, a *necessary* and *sufficient* condition for the existence of a unique solution is that

$$r(A) = n.$$

Let us consider the case where $m = n$, so that we can compare the above two results.

Suppose

$$r(A) = n.$$

Then

$$r(A\ \underset{\sim}{b}) \geqslant r(A) = n.$$

Since the columns of $(A\ \underset{\sim}{b})$ are elements of a vector space of dimension n, it follows that at most n of them are linearly independent, i.e.

$$r(A\ \underset{\sim}{b}) \leqslant n.$$

Hence

$$r(A\ \underset{\sim}{b}) = n.$$

That is,

$$r(A) = n \Rightarrow r(A) = r(A\ \underset{\sim}{b}).$$

On the other hand,

$$r(A) = r(A\ \underset{\sim}{b}) \not\Rightarrow r(A) = n.$$

Example 1

Example 1

The system

$$x - y = 2$$

$$x + y = 2$$

has the *unique* solution $x = 2$, $y = 0$.

Here

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, (A\, \underset{\sim}{b}) = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 1 & 2 \end{pmatrix}$$

and

$$r(A) = r(A\, \underset{\sim}{b}) = 2.$$

The system

$$2x + 2y = 2$$
$$x + y = 1$$

has *many* solutions of the form $x = \alpha$, $y = 1 - \alpha$.
In this case,

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}, \qquad (A\, \underset{\sim}{b}) = \begin{pmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

and

$$r(A) = r(A\, \underset{\sim}{b}) = 1 < 2.$$

■

Main Text
* * *

We have now completed, as far as we intend to go, our theoretical studies of existence and uniqueness. But before we go on to other things, we introduce a few terms and notation which are standard in the literature on linear algebra.

Definition 1
* *
Definition 2
* * *
Definition 3
* * *

When the number of equations is equal to the number of variables, the matrix of coefficients is said to be square. A matrix with m rows and n columns is often said to be of order $m \times n$. The square matrix A of order $n \times n$ (or sometimes "of order n") is said to be non-singular if $r(A) = n$. Otherwise A is called singular.

In general, the mapping

$$A : \underset{\sim}{x} \longmapsto A\underset{\sim}{x}$$

is a homomorphism, but if the matrix is non-singular, the mapping is an isomorphism of R^n to R^n. In this case, the mapping has an inverse which is also an isomorphism. We denote the matrix of the inverse mapping, and also the inverse mapping itself, by A^{-1}. In terms of mappings, we have

$$A^{-1} \circ A : \underset{\sim}{x} \longmapsto \underset{\sim}{x}$$

and

$$A \circ A^{-1} : \underset{\sim}{x} \longmapsto \underset{\sim}{x}.$$

If $I_{n,n}$ denotes the identity matrix of order n (see *Unit 23*, section 23.2.4), then in terms of matrices we have

$$A^{-1}A = A\,A^{-1} = I_{n,n}.$$

Definition 4
* * *

A^{-1} is called the inverse matrix of A.

If

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

has a unique solution, $\underset{\sim}{x}_p$, then we can write

$$\underset{\sim}{x}_p = A^{-1}\underset{\sim}{b}.$$

In section 26.3.2 and in *Unit 28, Linear Algebra IV*, we shall see how to calculate A^{-1}, and then this formula can prove useful; for instance, we may want the solutions of several sets of equations with the *same* A, but various $\underset{\sim}{b}$.

26.2.5 Summary

26.2.5

Summary
* * *

In section 26.2.1 we introduced the matrix form of a system of simultaneous linear equations:

$$A\underline{x} = \underline{b}.$$

In section 26.2.2 we gave a general discussion of the nature of the solution. In particular, we mentioned the *existence problem*:

Does a solution *exist*?

and the *uniqueness problem*:

Is there a *unique* solution?

In section 26.2.3 we discussed the existence problem in detail. We defined the *rank* of a matrix:

rank of A, $r(A)$ = (maximum number of linearly independent columns of A)

and the *augmented matrix*:

$$(A\ \underline{b}) = \left(A \;\vdots\; \underline{b}\right).$$

We gave the following theorem:

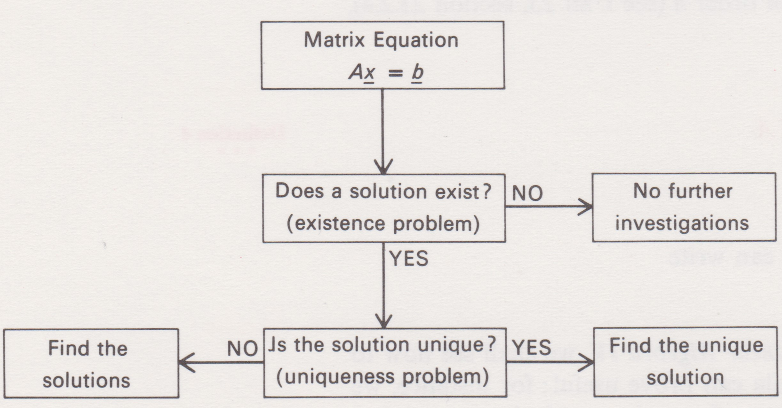
$$r(A) = r(A\ \underline{b})$$
$$\Rightarrow\ A\underline{x} = \underline{b} \text{ has a non-empty solution set.}$$

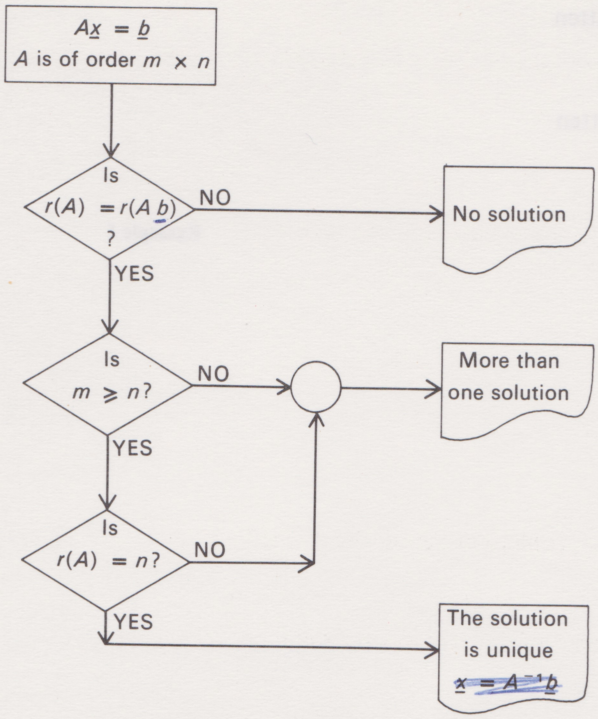
In section 26.2.4 we discussed the connection between the uniqueness problem and the rank of a matrix; we produced the following theorem:

If $A\underline{x} = \underline{b}$ is a system of linear equations in n unknowns, for which a solution exists, then

the solution is unique $\Leftrightarrow r(A) = n$.

A summary of the discussion on the existence and uniqueness problems is given below in diagrammatic form.





26.3 ELEMENTARY ROW OPERATIONS AND THEIR USE

26.3

26.3.1 Elementary Matrices

26.3.1

Introduction ★★

In section 26.1.2 we discussed one practical method for solving a system of equations, the Gauss elimination method. We then used the matrix notation, which is convenient for investigating the problems of the existence and uniqueness of solutions. We shall now look at the solution of a system of equations using the Gauss elimination method, but in terms of matrices. This has no practical advantage; in fact, it is a disadvantage. But it does allow us to discuss certain theoretical aspects of the method which have definite practical repercussions.

Main Text ★★★

In section 26.1.2 we defined three *elementary operations* used in the Gauss elimination method. We shall show that the equivalent operations, when the equations are written in matrix form, are multiplications of the coefficient matrix A by appropriate matrices. For simplicity we shall confine our attention to matrices of order 3×3, but the method we discuss is very general and can be applied to matrices of any order.

Definition 1 ★★★

We define three elementary row operations on a matrix A:

(i) interchange any two rows of the matrix;
(ii) multiply any row of the matrix by a non-zero number;
(iii) add a multiple of one row to another row.

We shall denote the first, second and third rows of the matrix A by R_1, R_2 and R_3 respectively.

Notation 1 ★★★

We denote an elementary row operation by E_i and abbreviate the descriptions as follows.

Interchange of R_1 and R_3 is written

$$E_1 : R_1 \longleftrightarrow R_3.$$

Multiplication of R_1 by the number k is written

$$E_2 : R_1 \longmapsto kR_1.$$

Addition of a multiple, k, of R_3 to R_1 is written

$$E_3 : R_1 \longmapsto R_1 + kR_3.$$

Example 1

Example 1

Consider

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & -2 \end{pmatrix}.$$

$E_1 : R_1 \longleftrightarrow R_2$ changes A to

$$E_1(A) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 3 \\ 1 & 0 & -2 \end{pmatrix}.$$

$E_2 : R_2 \longmapsto 3R_2$ changes A to

$$E_2(A) = \begin{pmatrix} 1 & 0 & 3 \\ 3 & 3 & 3 \\ 1 & 0 & -2 \end{pmatrix},$$

and changes $E_1(A)$ to

$$E_2(E_1(A)) = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & 9 \\ 1 & 0 & -2 \end{pmatrix}.$$

$E_3 : R_2 \longmapsto R_2 - 5R_3$ changes A to

$$\begin{pmatrix} 1 & 0 & 3 \\ -4 & 1 & 11 \\ 1 & 0 & -2 \end{pmatrix},$$

and changes $E_2(A)$ to

$$E_3(E_2(A)) = \begin{pmatrix} 1 & 0 & 3 \\ -2 & 3 & 13 \\ 1 & 0 & -2 \end{pmatrix}.$$

■

Discussion
* *

It is a remarkable fact that these operations can be performed on a matrix A by premultiplying A by particular matrices. The most direct method of demonstrating this is to find the matrices which perform the required operations. How do we find these matrices? There is a particularly simple way to do this. If we *assume* that such matrices exist, then they must perform the same operations on the identity matrix, in particular. For example, if E is *assumed* to be a matrix which interchanges the first and third rows of a 3×3 matrix, then

$$E = EI = E\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

since we know that E interchanges the first and third rows. Now let us premultiply a general matrix by E:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{pmatrix}.$$

We see that E has the required effect on *any* 3×3 matrix. This leads us to the following definition.

A matrix obtained from the unit matrix by an elementary row operation is called an elementary matrix.

Definition 2

Example 2

Example 2

We shall find the elementary matrices corresponding to the row operations

$$E_1 : R_1 \longleftrightarrow R_2,\ E_2 : R_2 \longmapsto 3R_2 \text{ and } E_3 : R_2 \longmapsto R_2 - 5R_3$$

which we used in Example 1. We shall then *premultiply* the matrix A of Example 1 by the elementary matrices found, and note the results obtained.

We begin with the identity matrix:

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$E_1 : R_1 \longleftrightarrow R_2$

Writing E_1 to stand for the matrix as well as the operation, we get

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$E_1A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 3 \\ 1 & 0 & -2 \end{pmatrix}.$$

$E_2 : R_2 \longmapsto 3R_2$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$E_2A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 \\ 3 & 3 & 3 \\ 1 & 0 & -2 \end{pmatrix}.$$

$E_3 : R_2 \longmapsto R_2 - 5R_3$

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix};$$

$$E_3A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 \\ -4 & 1 & 11 \\ 1 & 0 & -2 \end{pmatrix}.$$

You should compare these results with those of Example 1. ■

Exercise 1

Exercise 1
(2 minutes)

Find the elementary matrices of order 3×3 corresponding to each of the following:

(i) $R_2 \longleftrightarrow R_3$
(ii) $R_2 \longmapsto R_2 - 2R_1$
(iii) $R_3 \longmapsto R_3 + 2R_1 + 3R_2$
(iv) $R_2 \longmapsto R_2 - R_1 + 2R_3$

Verify that the elementary matrices have the desired effect by applying them to the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ -1 & -1 & -2 \end{pmatrix}.$$

■

Discussion
* * *

We have seen that, to each elementary row operation used in the Gauss elimination method, there corresponds an elementary matrix.

It appears that we have an isomorphism between

(1) the set of elementary operations carried out on the system of simultaneous equations, combined by successive performance,
and
(2) the set of elementary matrices, combined by matrix multiplication.

The object of the Gauss elimination method is to use a finite sequence of elementary operations to reduce a system of equations into any *equivalent* system which can be solved simply by back-substitution. We shall now reinterpret this in terms of matrices.

Suppose we start with the system of equations written in matrix form as

$$A\underset{\sim}{x} = \underset{\sim}{b},$$

that is,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

The object is to reduce this system to the equivalent system

$$C\underset{\sim}{x} = \underset{\sim}{d},$$

that is,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

If E_1 represents the first elementary operation which we carry out on the original system, we now premultiply both sides of the matrix equation by E_1, to obtain

$$E_1A\underset{\sim}{x} = E_1\underset{\sim}{b}.$$

This new equation represents a system of equations equivalent to the first. Notice that by premultiplying A and $\underset{\sim}{b}$ by E_1 we are effectively doing the elementary operation. So there is no point, in any practical calculation, in finding the matrix E and actually doing the matrix multiplication: it is much easier to carry out the appropriate row operation on the augmented matrix $(A\ \underset{\sim}{b})$. The *only* value in knowing of the existence of the matrix E is in theoretical considerations which may have practical consequences, but the matrix E itself is not used in practice.

We have seen that each elementary operation corresponds to an elementary matrix. It follows that a sequence of elementary operations corresponds to a product of elementary matrices. Suppose the sequence of matrices, in order of usage, is $E_1, E_2, E_3, \ldots, E_s$. Then, what we are doing can be written in the form

$$(E_s \ldots E_3 E_2 E_1 A)\underset{\sim}{x} = E_s \ldots E_3 E_2 E_1 \underset{\sim}{b}.$$

Notice that, because we are always premultiplying by the E's, the column vector $\underset{\sim}{x}$ is never involved. When setting out numerical calculations we drop the $\underset{\sim}{x}$ and just keep the augmented matrix $(A\ \underset{\sim}{b})$ and manipulate that. We have

$$\begin{aligned}(C\ \underset{\sim}{d}) &= (E_s \ldots E_3 E_2 E_1)(A\ \underset{\sim}{b})\\ &= P(A\ \underset{\sim}{b}),\end{aligned}$$

where P is the matrix obtained by multiplying all the E's together. Incidentally, we know that *if* we can get from $(A\ \underset{\sim}{b})$ to $(C\ \underset{\sim}{d})$, we can reverse each step (each elementary operation can be reversed by another elementary operation of the same kind) and get from $(C\ \underset{\sim}{d})$ back to $(A\ \underset{\sim}{b})$. This means that the matrix P is non-singular, i.e. it has an inverse matrix P^{-1}. Notice that we said "*if* we can get from $(A\ \underset{\sim}{b})$ to $(C\ \underset{\sim}{d})$": we have no guarantee that we can. In fact, it is always possible, although some of the c's may be zero. If the solution of the system of equations is unique, i.e. if $r(A) = 3$ (or, in general, if $r(A) = n$ for a square $n \times n$ matrix), then none of the c's on the leading diagonal is zero.

We have now achieved the object of this section, which was to interpret the Gauss elimination method in terms of matrices. There are a number of interesting consequences and refinements which we shall consider in the remainder of this text.

Exercise 2

Exercise 2
(2 minutes)

Find elementary matrices which reduce the system

$$\begin{pmatrix} 2 & 1 & -1 \\ 6 & -1 & -9 \\ 4 & 3 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix}$$

to the system

$$\begin{pmatrix} 2 & 1 & -1 \\ 0 & -4 & -6 \\ 0 & 0 & \frac{3}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ -\frac{3}{2} \end{pmatrix}.$$

Hence find the matrix P and verify that

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 3 \\ 0 & -4 & -6 & -2 \\ 0 & 0 & \frac{3}{2} & -\frac{3}{2} \end{array}\right) = P\left(\begin{array}{ccc|c} 2 & 1 & -1 & 3 \\ 6 & -1 & -9 & 7 \\ 4 & 3 & 1 & 5 \end{array}\right) \quad \blacksquare$$

Solution 1

Solution 1

(i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(iii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix}$ (iv) $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ ■

Solution 2

Solution 2

The E's are not unique: they will depend on which elementary operations are chosen. But P is unique:

$$\begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -\frac{11}{4} & \frac{1}{4} & 1 \end{pmatrix}$$

A possible choice of elementary matrices is

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{4} & 1 \end{pmatrix}}_{E_3} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}}_{E_2} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{E_1} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -\frac{11}{4} & \frac{1}{4} & 1 \end{pmatrix}}_{P}.$$

■

26.3.2 The Inverse of a Matrix

26.3.2

Main Text * *

There are a number of ways of finding the inverse of a non-singular matrix A. In this section we shall exploit the elementary matrix technique discussed in the last section. This is a good example of the way in which elementary matrices, although not themselves practical, lead to practical methods.

Suppose that A is a non-singular matrix and that X is its inverse. Then

$$XA = AX = I,$$

where A, X and I are square matrices of the same order. We can consider the equation $AX = I$ as an equation from which we wish to determine the unknown matrix X. It is, in fact, not very different from our previous equation. For instance, if we suppose that A, X and I are all 3×3, then we can write

$$A(\underline{x}_1 \quad \underline{x}_2 \quad \underline{x}_3) = (\underline{i}_1 \quad \underline{i}_2 \quad \underline{i}_3),$$

where we have expressed the matrices X and I in terms of their column vectors in the usual way. This *one* matrix equation is then equivalent to the *three* matrix equations

$$A\underline{x}_1 = \underline{i}_1, \qquad A\underline{x}_2 = \underline{i}_2, \qquad A\underline{x}_3 = \underline{i}_3,$$

and we are back to the problem of solving the equation $A\underline{x} = \underline{b}$, except that we now have to solve three matrix equations instead of one.

This suggests that the same approach might help here. We can perhaps find a sequence of elementary operations (with corresponding elementary matrices) which together form the Gauss elimination method. The same sequence of operations would, of course, do for all the three equations. This approach would certainly work, but a little more effort will

give a much bigger return. The Gauss elimination method requires back-substitution; we can avoid back-substitution if we carry out some more elementary operations. Instead of reducing A (in the 3×3 case) to the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{pmatrix}, \quad \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{pmatrix}$$

we could carry on, and try to reduce it all the way down to

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

still using elementary operations. Then we could read off the solution without back-substitution.

Let us suppose that we can find a sequence of elementary operations which transform A into I. We shall denote the corresponding elementary matrices by $E_1, E_2, \ldots, E_s$. Then starting from our original equation

$$AX = I,$$

we have

$$(E_s \ldots E_2E_1A)X = E_s \ldots E_2E_1I,$$

that is,

$$IX = E_s \ldots E_2E_1I,$$

or

$$X = E_s \ldots E_2E_1I.$$

This is a remarkable result which has practical consequences. It tells us that if we can find a sequence of elementary operations which transforms A into I, that same sequence of operations will transform I into the inverse of A. The elementary matrices give us the justification, but in practice we do not use these — we use their effects — the elementary operations.

Example 1

Example 1

Find the inverse of the matrix

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 6 & -1 & -9 \\ 4 & 3 & 1 \end{pmatrix}.$$

There is no harm in assuming that the inverse exists: we have no reason to suppose either that it does or that it does not, but the proof of the pudding will be in the eating. If we can find a sequence of elementary operations which transforms A into I, then the inverse of A exists. Since we are going to perform the same elementary operations on the rows of A and I, we get ourselves into battle array, dropping matrix brackets.

$$\begin{array}{rrr|rrr|r} 2 & 1 & -1 & 1 & 0 & 0 & 3 \\ 6 & -1 & -9 & 0 & 1 & 0 & -3 \\ 4 & 3 & 1 & 0 & 0 & 1 & 9 \end{array}$$

You may be wondering where the last column came from. As we mentioned in our notes on the Gauss elimination method, any good numerical method should incorporate a check of some sort. So we have introduced an extra figure at the end of each row, which is the sum of the numbers in that row. In the calculation we treat it as part of that row, so that whatever we do to the row (by an elementary operation), the last number

should still be the sum of the numbers in that row. If, after a certain step, we find that the sum-check fails, then there is an error somewhere in that step. As long as we remember to check the row sum after each step, we should be *reasonably* sure of getting the whole calculation right. (Only *reasonably* sure, because we may have made compensating errors.) We now proceed *systematically* to produce the matrix I in the first three columns, indicating each step by our usual notation.

$R_1 \longmapsto \frac{1}{2}R_1$

$$\begin{array}{ccc|ccc|c} 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{3}{2} \\ 6 & -1 & -9 & 0 & 1 & 0 & -3 \\ 4 & 3 & 1 & 0 & 0 & 1 & 9 \end{array}$$

(When you get adept at the game, then you don't need to copy down *all* the rows at each step, but only the ones that change; for example, the first row above. We shall not do this because it can be a bit confusing for the beginner, but we shall occasionally do more than one step at one go.)

$R_2 \longmapsto R_2 - 6R_1\,;\; R_3 \longmapsto R_3 - 4R_1$

$$\begin{array}{ccc|ccc|c} 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{3}{2} \\ 0 & -4 & -6 & -3 & 1 & 0 & -12 \\ 0 & 1 & 3 & -2 & 0 & 1 & 3 \end{array}$$

We now have one column correct.

$R_2 \longmapsto -\frac{1}{4}R_2$

$$\begin{array}{ccc|ccc|c} 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{3}{2} \\ 0 & 1 & \frac{3}{2} & \frac{3}{4} & -\frac{1}{4} & 0 & 3 \\ 0 & 1 & 3 & -2 & 0 & 1 & 3 \end{array}$$

$R_1 \longmapsto R_1 - \frac{1}{2}R_2\,;\; R_3 \longmapsto R_3 - R_2$

$$\begin{array}{ccc|ccc|c} 1 & 0 & -\frac{5}{4} & \frac{1}{8} & \frac{1}{8} & 0 & 0 \\ 0 & 1 & \frac{3}{2} & \frac{3}{4} & -\frac{1}{4} & 0 & 3 \\ 0 & 0 & \frac{3}{2} & -\frac{11}{4} & \frac{1}{4} & 1 & 0 \end{array}$$

We now have two columns correct.

$R_3 \longmapsto \frac{2}{3}R_3$

$$\begin{array}{ccc|ccc|c} 1 & 0 & -\frac{5}{4} & \frac{1}{8} & \frac{1}{8} & 0 & 0 \\ 0 & 1 & \frac{3}{2} & \frac{3}{4} & -\frac{1}{4} & 0 & 3 \\ 0 & 0 & 1 & -\frac{11}{6} & \frac{1}{6} & \frac{2}{3} & 0 \end{array}$$

$R_1 \longmapsto R_1 + \frac{5}{4}R_3\,;\; R_2 \longmapsto R_2 - \frac{3}{2}R_3$

$$\begin{array}{ccc|ccc|c} 1 & 0 & 0 & -\frac{13}{6} & \frac{1}{3} & \frac{5}{6} & 0 \\ 0 & 1 & 0 & \frac{7}{2} & -\frac{1}{2} & -1 & 3 \\ 0 & 0 & 1 & -\frac{11}{6} & \frac{1}{6} & \frac{2}{3} & 0 \end{array}$$

We now have all three columns correct.

If our theory is correct, then the inverse of A is

$$\begin{pmatrix} -\frac{13}{6} & \frac{1}{3} & \frac{5}{6} \\ \frac{7}{2} & -\frac{1}{2} & -1 \\ -\frac{11}{6} & \frac{1}{6} & \frac{2}{3} \end{pmatrix}.$$

To check that this matrix is indeed the inverse of the matrix A, do Exercise 1.

Exercise 1
(3 minutes)

Exercise 1

Given that

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 6 & -1 & -9 \\ 4 & 3 & 1 \end{pmatrix} \text{ and } X = \begin{pmatrix} -\frac{13}{6} & \frac{1}{3} & \frac{5}{6} \\ \frac{7}{2} & -\frac{1}{2} & -1 \\ -\frac{11}{6} & \frac{1}{6} & \frac{2}{3} \end{pmatrix},$$

show that

$$XA = I = AX,$$

where

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

■

Exercise 2
(3 minutes)

Exercise 2

Find the inverse matrices of

$$\text{(i)} \begin{pmatrix} 1 & -1 & 0 \\ 3 & 7 & 1 \\ 2 & -14 & 2 \end{pmatrix} \quad \text{(ii)} \begin{pmatrix} 0 & 1 & -1 \\ 3 & 6 & 2 \\ -1 & 0 & 5 \end{pmatrix}.$$

Set out the calculations by hand, or, if you have the time and inclination, use the computer program (given in the Appendix) to perform the elementary operations for you. ■

Discussion
* *

We now have a technique for finding the inverse of any non-singular matrix. The question may naturally arise in your mind: What is the use of it? We have briefly answered this question earlier, but we shall now discuss it more fully.

If we have a system of equations

$$A\underset{\sim}{x} = \underset{\sim}{b},$$

where A is a non-singular square matrix, then, since the inverse matrix, A^{-1}, represents the inverse isomorphism, there is a unique solution to the system; it is

$$\underset{\sim}{x} = A^{-1}\underset{\sim}{b}.$$

The important thing about it is that, once we know A^{-1}, we can find a unique solution for *any* column vector $\underset{\sim}{b}$. This is a distinct advantage over the Gauss elimination method, where, for a different choice of $\underset{\sim}{b}$, we would have to do the Gauss elimination again. (Although, of course, only the $\underset{\sim}{b}$ column would be different and so, provided we had kept a record of our previous calculations, the work involved would not be as hard as it may seem at first sight.)

We shall see in *Unit 28, Linear Algebra IV* that if we have a set of systems of equations

$$A\underset{\sim}{x} = \underset{\sim}{b}_i \quad (i = 1, 2, \ldots, n)$$

with the same coefficient matrix but a number of different right-hand sides, then, provided $n > 3$, it is well worth while finding A^{-1}, and for each i, calculating the solution

$$A^{-1}\underset{\sim}{b}_i.$$

Although we have restricted our considerations here to non-singular matrices, and we have no effective method of looking at a square matrix and telling whether it is non-singular, the method described in this section

(*continued on page 38*)

Solution 1

Solution 1

Just multiply out. ■

Solution 2

Solution 2

(i)

$$\text{The inverse matrix of } \begin{pmatrix} 1 & -1 & 0 \\ 3 & 7 & 1 \\ 2 & -14 & 2 \end{pmatrix}$$

$$\text{is } \begin{pmatrix} \frac{7}{8} & \frac{1}{16} & -\frac{1}{32} \\ -\frac{1}{8} & \frac{1}{16} & -\frac{1}{32} \\ -\frac{7}{4} & \frac{3}{8} & \frac{5}{16} \end{pmatrix}.$$

If you did it by using the computer program, you should finish up with

$$\begin{pmatrix} .875 & .0625 & -.03125 \\ -.125 & .0625 & -.03125 \\ -1.75 & .375 & .3125 \end{pmatrix}.$$

(ii)

$$\text{The inverse matrix of } \begin{pmatrix} 0 & 1 & -1 \\ 3 & 6 & 2 \\ -1 & 0 & 5 \end{pmatrix}$$

$$\text{is } \begin{pmatrix} -\frac{30}{23} & \frac{5}{23} & -\frac{8}{23} \\ \frac{17}{23} & \frac{1}{23} & \frac{3}{23} \\ -\frac{6}{23} & \frac{1}{23} & \frac{3}{23} \end{pmatrix}.$$

If you did this one by using the computer program, you should get something like

$$\begin{pmatrix} -1.30435 & .217391 & -.347826 \\ .73913 & 4.34783\text{E-}02 & .130435 \\ -.26087 & 4.34783\text{E-}02 & .130435 \end{pmatrix}.$$

■

(*continued from page 37*)

is still useful. Even if it does not lead to the inverse matrix, some of the same steps will lead to a solution (or solutions) if one exists: if no solution exists, it will tell us this as well. We shall not go into these points here. They are not difficult, but we have gone far enough into the theory of linear equations for the time being.

We shall turn briefly, in the next section, to the calculation of rank. This involves a similar discussion to the one above.

26.3.3 Calculation of the Rank of a Matrix

26.3.3

Main Text
* *

Although we have stated some of the results in this text in terms of the rank of a matrix, we have not discussed an effective way of calculating it (our examples were artificially easy). In this section we shall make good this deficiency. We shall restrict our discussion to square matrices for simplicity.

Given a square matrix A, we have defined the rank of A as the maximum number of linearly independent column vectors in A.

If A is an $n \times n$ matrix, then we can associate with A the mapping

$$A : \underset{\sim}{x} \longmapsto A\underset{\sim}{x} \qquad (\underset{\sim}{x} \in R^n),$$

and we have shown in section 26.2.4 that

$$r(A) = \text{dimension of } A(R^n).$$

We shall consider the effect of one of the elementary row operations on the rank of A. Although we cannot necessarily readily tell the rank of A, we can tell the rank of a much simplified associated matrix (for instance, of the form obtained in the Gauss elimination process).

Let E be the matrix corresponding to one of the elementary row operations. The corresponding mapping

$$E : \underset{\sim}{x} \longmapsto E\underset{\sim}{x} \qquad (\underset{\sim}{x} \in R^n)$$

is one-one, because, as we have already noted, every elementary row operation has an inverse operation of the same type.

Consider the composite mapping

$$E \circ A : \underset{\sim}{x} \longmapsto EA\underset{\sim}{x} \qquad (x \in R^n)$$

with matrix EA. We have

$$\begin{aligned} r(EA) &= \text{dimension of } E \circ A(R^n) \\ &= \text{dimension of } E(A(R^n)) \end{aligned}$$

But, as we saw in section 26.2.4, since E is one-one, the kernel of the E mapping contains just the zero element, and so is of dimension zero; that is,

$$\text{dimension of } E(A(R^n)) = \text{dimension of } A(R^n) = r(A).$$

So

$$r(A) = r(EA).$$

This means that the rank of A is *unaffected by an elementary row operation.* This leads to a practical method of calculating rank.

Example 1

Example 1

We calculate the rank of the matrix

$$A = \begin{pmatrix} 1 & 0 & -1 \\ -2 & 1 & 3 \\ 4 & 2 & -2 \end{pmatrix}$$

by *systematically* reducing the elements below the leading diagonal (from top left to bottom right) to zero, using elementary row operations. We use a row sum, as described in the previous section, to act as a check for the arithmetic.

$$\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ -2 & 1 & 3 & 2 \\ 4 & 2 & -2 & 4 \end{array}$$

$R_2 \longmapsto R_2 + 2R_1\,;\; R_3 \longmapsto R_3 - 4R_1$

$$\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 2 & 2 & 4 \end{array}$$

$R_3 \longmapsto R_3 - 2R_2$

$$\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}$$

It is now easy to see that the first two columns are linearly independent, but

$$\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = -1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

so $r(A) = 2$.

In fact, if we use the result, mentioned in section 26.2.3, that the row rank (the maximum number of linearly independent row vectors) is equal to the column rank, we can see the rank of A even more easily, since the last row consists entirely of zeros. In general, if we carry out this reduction, the rank of A is equal to the number of "non-zero" rows in the reduced matrix. ■

Exercise 1

Exercise 1
(2 minutes)

Calculate the rank of the following matrix:

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & -1 & -1 \end{pmatrix}.$$

■

Exercise 2

Exercise 2
(4 minutes)

Use methods similar to those in the text to prove that, if B is an $n \times n$ non-singular matrix and A is any $n \times n$ matrix, then

$$r(BA) = r(A).$$

■

Exercise 3

Exercise 3
(5 minutes)

If A and B are any $n \times n$ matrices, prove that

$$r(BA) \leqslant r(A).$$

■

26.4 SUMMARY

26.4

Summary
★ ★

We have come full circle. Towards the beginning of the text we described the Gauss elimination method for solving a set of simultaneous equations in terms of elementary operations on the equations in the system.

Then we discussed the problem of the existence and uniqueness of the solution of such a system and obtained two theoretical results in terms of the ranks of the matrices describing the system. These results were:

The system of linear equations represented by

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

has a non-empty solution set if $r(A) = r(A \; \underset{\sim}{b})$.

The system of n equations in n unknowns represented by

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

has a unique solution if and only if

$$r(A) = n.$$

We then "changed direction" and defined elementary matrices. These correspond to the elementary operations in the Gauss elimination method. We thus obtained a matrix representation of this method. Although this had no direct practical interest, it had practical consequences. It allowed us to develop:

(i) a method for finding the inverse of a non-singular matrix,
and
(ii) a method for calculating the rank.

Since both these methods use essentially the same systematic procedure as the Gauss elimination method, we can use the same calculations to determine:

(i) the solution to a system of equations,
(ii) the inverse of the matrix of coefficients, if it exists,
(iii) the rank of the matrix of coefficients or the augmented matrix,

whichever is of interest.

Solution 1

Solution 1

The rank of the matrix is 3. ■

Solution 2

Solution 2

Since B is non-singular, we can apply the same argument to B as to E in the text preceding Example 1. That is,

$$r(BA) = \text{dimension of } B(A(R^n)) = \text{dimension of } A(R^n) = r(A).$$ ■

Solution 3

Solution 3

$$r(BA) = \text{dimension of } B(A(R^n)),$$

and from the dimension theorem,

$$\text{dimension of } B(A(R^n)) = \text{dimension of } A(R^n) - \text{dimension of kernel of } B \text{ with domain } A(R^n).$$

Since the dimension of any vector space is greater than or equal to zero,

$$r(BA) = \text{dimension of } B(A(R^n)) \leqslant \text{dimension of } A(R^n) = r(A).$$

Similarly, $r(BA) \leqslant r(B)$ as well; so the rank of the product of two matrices is less than or equal to the rank of either matrix. ■

26.5 APPENDIX

26.5

Computer Program $ROWOP

Appendix

This program will perform the row operations which you need in order to simplify or invert a matrix. In particular it will:

(a) Interchange two rows, upon the instruction INT, followed by the numbers of the two rows you wish to interchange.

(b) Multiply a row by any non-zero number, upon the instruction MULT, followed by the number of the row, then the number you wish to multiply it by.

(c) Multiply one row by any number, and add the result to another row. It does this on the instruction MARS (short for "Multiply and Add RowS"), followed by

- (i) the number of the row to be multiplied;
- (ii) the number by which it is to be multiplied;
- (iii) the row to which the result is to be added.

Sometimes the number you want to multiply by is the reciprocal of one of the entries of the matrix. To cope with this eventuality, we have also prepared the program so that the computer will accept the following instructions (logically redundant, but practically useful):

(d) Divide a row by any non-zero number, upon the instruction DIV, followed by the number of the row, then the number you wish to divide it by.

(e) Divide one row by any non-zero number, and add the result to another row. It does this on the instruction DARS (short for "Divide and Add RowS") followed by

- (i) the number of the row to be divided;
- (ii) the number by which it is to be divided;
- (iii) the row to which the result is to be added.

Running Instructions

Log in at the computer terminal in the normal way. When you have logged in, type

GET-$ROWOP

followed by the carriage-return, then type

RUN

followed by the carriage-return again. The computer will print

?

and you must input two numbers (which must not be greater than 10); these are the number of rows, and the number of columns, of the matrix you are going to work with. For instance, if you type

4,3

followed by the carriage-return, then you have told the computer to expect a 4×3 matrix, i.e. a matrix with 4 rows and 3 columns.

The computer will again print

?

and now you must input the entries of the matrix, in lexicographic order, i.e. the entries of the first row, in order, then those of the second row, in order, etc. For instance, if you want to input the matrix

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 1 & 3 \\ 0 & 4 & 4 \end{pmatrix}$$

(which is a 4×3 matrix), your input at this juncture would be

1,1,2,1,2,3,2,1,3,0,4,4

followed by the carriage-return.

The computer will again print

?

and this time you can give one of the instructions INT, MULT, MARS, DIV or DARS, followed by the carriage-return. You will again get

?

and you must put in either two or three numbers, as already indicated, to complete the specification of the row operation. For instance, suppose you had typed DARS and wished to divide row 3 by 2 and then subtract it from row 1. This is equivalent to dividing row 3 by -2 and adding it to row 1, so you type

3, -2, 1

followed by the carriage-return, and the computer will perform the operation and type

?

awaiting further instructions.

At this or any other stage (even immediately after inputting the matrix, if you wish), you can, instead of giving an INT, MULT, MARS, DIV or DARS instruction, give the instructions PRINT, RET, NEXT MATRIX or FIN. The PRINT instruction will cause the computer to print out the

matrix as it has evolved so far. If you wish to undo the effect of the operation you have just performed, the instruction RET will regain the matrix as it was prior to the last row operation you performed.

When you have finished working on a matrix and wish to go on to the next one, give the instruction NEXT MATRIX, and the computer will go back to the beginning of the program, ready for you to input the details of another matrix.

When you have finished with all the matrices, give the instruction FIN. The computer will terminate the program, and print

DONE

You sign off at this point with

BYE

and record the time taken.

Error Diagnosis

If at any stage in the program, you type in something which does not make sense at that particular point, the computer will print out some statement which will tell you what was wrong. Here is a (hopefully) complete list of what to expect.

If you get

EXTRA INPUT — WARNING ONLY

this means that you tried to put in too many numbers. For instance, after a DIV instruction you may have typed 1, 2, 3 by mistake. The computer will take the 1 and 2 and act on them, and the program will proceed.

If you get

??

then you did not put in enough numbers. For instance, after a MULT instruction you may have typed 1, 2. In this case, you must type in the missing number(s), and the program will proceed.

If you get

???

this means that you tried to put in something other than numbers (e.g. MULT, or something), where you were supposed to give numbers. In this case you must give the correct response, and the program will proceed.

The above messages are part of the computer's automatic repertoire. The messages which follow are part of the ROWOP program, and subject to our control. We have arranged it so that, if you make three of the following types of error *in succession*, the program will terminate. You will have to type

BYE

and record the time taken, then go away and re-think what you are doing.

If when specifying the number of rows and columns of the matrix, you fail to give positive integers less than or equal to 10, you will get the message

INADMISSIBLE VALUES

and you will have to specify the values again.

If when giving numbers to specify row operations, you give numbers which are not positive integers less than or equal to the number of rows in the matrix, you will again get

INADMISSIBLE VALUES

and this time you must go right back and re-specify an operation, with a command such as INT, MULT, etc.

If when you are supposed to give a command such as INT, MULT, etc., you fail to do so, you will get the message

INADMISSIBLE OPERATIONAL COMMAND

and you must specify such a command properly.

If you specify a row operation whose effect would be to reduce a row to zero or try to divide by zero you will get

LETHAL ROW OPERATION

The computer will not carry out this operation, and will take you back to the point in the program where you have to specify INT, MULT, etc., again.

Summary

Having logged in, recorded the time, typed GET–$ROWOP and RUN, you must input the number of rows and columns of the matrix. Then you must input the entries of the matrix. Then you have, at each step, the following choice of instructions:

INT

MULT

MARS

DIV

DARS

RET

PRINT

NEXT MATRIX

FIN

After INT, input *i*,*j*, and the computer interchanges rows *i* and *j*.

After MULT, input *i*,*x*, and the computer multiplies row *i* by *x*.

After MARS, input *i*,*x*,*j*, and the computer multiplies row *i* by *x* and adds the result to row *j*.

After DIV, input *i*,*x*, and the computer divides row *i* by *x*.

After DARS, input *i*,*x*,*j*, and the computer divides row *i* by *x* and adds the result to row *j*.

The instruction RET undoes the last row operation that was performed.

The instruction PRINT prints out the matrix so far.

The instruction NEXT MATRIX prepares the computer to receive the next matrix.

The instruction FIN causes the program to terminate. The computer types DONE, you type BYE and record the time.

Unit No.		Title of Text
1		Functions
2		Errors and Accuracy
3		Operations and Morphisms
4		Finite Differences
5	NO TEXT	
6		Inequalities
7		Sequences and Limits I
8		Computing I
9		Integration I
10	NO TEXT	
11		Logic I — Boolean Algebra
12		Differentiation I
13		Integration II
14		Sequences and Limits II
15		Differentiation II
16		Probability and Statistics I
17		Logic II — Proof
18		Probability and Statistics II
19		Relations
20		Computing II
21		Probability and Statistics III
22		Linear Algebra I
23		Linear Algebra II
24		Differential Equations I
25	NO TEXT	
26		Linear Algebra III
27		Complex Numbers I
28		Linear Algebra IV
29		Complex Numbers II
30		Groups I
31		Differential Equations II
32	NO TEXT	
33		Groups II
34		Number Systems
35		Topology
36		Mathematical Structures